Spring 1995

A Feasibility Study of Dynamical Assimilation of Tide Gauge Data in the Chesapeake Bay

Yvette H. Spitz
Old Dominion University

A FEASIBILITY STUDY OF DYNAMICAL ASSIMILATION OF TIDE GAUGE DATA IN THE CHESAPEAKE BAY

by

Yvette H. Spitz

M.Sc., April 1990, Florida State University, Tallahassee, Florida, U.S.A.
Licence in Physics, October 1983, Liège University, Liège, Belgium

A Dissertation submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

Oceanography

Old Dominion University
May, 1995

Approved by:

John M. Klinck (Director)

Larry P. Atkinson

Gabriel T. Csanady

Linda M. Lawson

Ionel M. Navon

Abstract

A FEASIBILITY STUDY OF DYNAMICAL ASSIMILATION OF TIDE GAUGE DATA IN THE CHESAPEAKE BAY

Yvette Huberte Spitz

Old Dominion University, 1995
Director: Dr. J. M. Klinck

The feasibility of dynamical assimilation of surface elevation from tide gauges is investigated to estimate the bottom drag coefficient and surface stress as a first step in improving modeled tidal and wind-driven circulation in the Chesapeake Bay. A two-dimensional shallow water model and an adjoint variational method with a limited memory quasi-Newton optimization algorithm are used to achieve this goal.

Assimilation of tide gauge observations from ten permanent stations in the Bay and use of a two-dimensional model adequately estimate the bottom drag coefficient, wind stress and surface elevation at the Bay mouth. Subsequent use of these estimates in the circulation model considerably improves the modeled surface elevation in the entire Bay. Assimilation of predicted tidal elevations yields a drag coefficient, defined in the hydraulic way, varying between 2.5×10^{-4} and 3.1×10^{-3}. The bottom drag coefficient displays a periodicity corresponding to the spring-neap tide cycle. From assimilation of actual tide gauge observations, it is found that the fortnightly modulation is altered during frontal passage. Furthermore, the response of the sea surface to the wind forcing is found to be more important in the lower Bay than in the upper Bay, where the barometric pressure effect could be more important.

In addition, identical twin experiments with model generated data show that a

penalty term has to be added to the simple cost function defined as the distance between modeled and observed surface elevation in order to assure smoothness of the surface elevation field at the Bay mouth. Classical scaling of the parameters to bring them to the same order of magnitude was not effective in accelerating the convergence during the assimilation procedure and yielded larger errors in the estimated parameters.

Acknowledgments

I wish to express my gratitude to Dr. John Klinck, my dissertation advisor, for his support, guidance and encouragement in pursuing the independent research which comprises this work. I thank the members of my committee, Drs. Larry Atkinson, Gabriel Csanady, Linda Lawson and Ionel Navon, for their interest in my work as well as their help and suggestions. My special thanks go to Dr. Linda Lawson for her great help, expertise in data assimilation, kindness, constant encouragement and for the understanding that she has shown these past three years. I am also grateful to Dr. Ionel Navon and Dr. X. Zou of Florida State University for sharing their expertise in data assimilation as well as in adjoint model coding and pointing out the important references related to the variational method. I also thank Dr. Larry Atkinson for his support and for providing a wonderful working environment at CCPO. I thank Dr. Eileen Hofmann, CCPO, for giving me the opportunity to apply the variational data assimilation technique to a marine ecosystem model. This work with Drs. Linda Lawson, ETSU, and Eileen Hofmann afforded me the opportunity to explore the field of biological-physical modeling. Thanks are due to Dr. Arnoldo Valle-Levinson, CCPO, for the numerous productive discussions we had regarding tidal modeling and the Chesapeake Bay. I thank the Commonwealth of Virginia and the Center for Coastal Physical Oceanography for supporting this research.

I thank the Management Unit of the Mathematical Models of the North Sea and Scheldt Estuary, especially Dr. Georges Pichot and José Ozer, for allowing me to use their model. Dr. Steve Gill and Captain Carl Fisher, of NOAA, provided the data and tidal analysis used in this work. Dr. John Hamrick, VIMS, provided the model grid used in the simulations. Their assistance is greatly appreciated.

Finally, I am very grateful to my mother and sister who have always been supportive during my studies far from home. All the faculty members, students and staff members at CCPO and from the Department of Oceanography are thanked for their friendship and encouragement over the last four years.

Contents

List of Tables

List of Figures

1 Introduction

Chesapeake Bay, the largest estuary in the United States, is not only a major waterway for commercial marine transportation, naval operations and recreational boating but also a highly productive marine environment. For instance, larvae and postlarvae of fishes and crab, which have been spawned in the coastal ocean, re-enter the Bay in late summer and early fall. Because of its interrelation with other processes taking place in the Chesapeake Bay, *e.g.*, water quality and biological productivity, circulation in the Bay is probably the first process that needs to be understood.

The main components of the estuarine circulation are the tidal, gravitational and wind-driven circulation. During the last 40 years, circulation in the Bay has mainly been studied from observations of temperature, salinity, sea level, and currents. Pritchard (Officer, 1976) has extensively studied the gravitational circulation while Wang and Elliott (1978), Wang (1979a,b) have analyzed the response of the Bay to the wind forcing, and recently, Paraso and Valle-Levinson (1995) have studied the response of the lower Bay to atmospheric forcings, *i.e.*, wind and barometric pressure. An extensive analysis of the tidal circulation from tide gauge sea level and current measurements was done by Fisher (1986). These studies helped to recognize that wind and bottom stress greatly influence the circulation in the Bay. However, wind stress and bottom friction are difficult to estimate. For instance, the wind speed and direction are essentially measured at major airports on the western side of the Bay. But, conversion of the wind on land to wind over water is not an easy task. For instance, Goodrich (1985) showed that a different correction has to be done

to the alongshore and cross-shore wind components measured at airports. Bottom friction, while hard to measure, is usually defined in the two-dimensional models as a quadratic function of the vertically-integrated velocity. An empirical parameter, the bottom drag coefficient, is adjusted for a best fit between the modeled surface elevations and observations at tide gauge stations (Crean *et al.*, 1988).

Recently, variational data assimilation and inverse methods have been used to determine the bottom drag coefficient. Using a two-dimensional model and assimilation of tide gauges data, Das and Lardner (1992), Lardner *et al.* (1993) showed that a depth correction and bottom friction coefficient can be estimated. Using an inverse method and tidal current measurements, Bang (1994) estimated the bottom drag coefficient in the Chesapeake Bay. The study from Lardner *et al.* (1993) in the Arabian Gulf is probably most closely related to our study. However, several major differences can be pointed out. In our study, the bottom drag coefficient is not only estimated but also the surface forcing, and the tide gauges located close to the coast are used, which is far more challenging than the use of open water gauges. Finally, the drag coefficient parameterization is different in both studies.

Data assimilation techniques have been recently developed in meteorology as well as in oceanography and are numerous (Ghil *et al.*, 1981; Navon, 1986; Ghil and Malanotte-Rizzoli, 1991; Navon *et al.*, 1992b,c). Most of these methods fit in one of the following classes: i) local polynomial interpolation methods (Cressman, 1959), ii) statistical (optimal) interpolation methods (Lorenc, 1981), iii) variational numerical analysis. The latter technique was originated in meteorology by Sasaki (1955, 1970) and has been developed considerably since then. It addresses the question of sensitivity analysis (Cacuci, 1981; Hall *et al.*, 1982; Hall and Cacuci, 1983; Cacuci and Hall, 1984; Cacuci, 1988; Zou *et al.*, 1993b), variational adjustment (Lewis and Derber, 1985; Talagrand and Courtier, 1987; Thacker and Long, 1988, Navon *et al.*, 1992a) and parameter estimation (Panchang and O'Brien, 1989; Das

and Lardner, 1991; Smedstad and O'Brien, 1991; Yu and O'Brien, 1991; Zou *et al.*, 1992b; Lardner, 1993; Lawson *et al.*, 1995a,b). This list of references represents only a small sample of what has been done in data assimilation.

The objective of variational data assimilation is to minimize a cost function with respect to the control variables by minimizing the misfit between model equivalents of the data and observations. The goodness-of-fit of the model equivalents of the data to the observations is measured by a cost function, which is minimized by adjusting the control variables. Most of the optimization algorithms are based on iterative descent large-scale unconstrained local minimization methods which require the computation of the gradient of the cost function with respect to the control variables. The adjoint of the model equations is used to compute the gradient of the cost function. The adjoint model equations can be derived by using different methods: the derivation of Euler-Lagrange equations (Morse and Feshbach, 1953), the control theory (Le Dimet and Talagrand, 1986), the Lagrange multiplier approach (Thacker and Long, 1988). In those methods, the continuous adjoint equations are first derived then discretized. However, recent studies have shown that the adjoint model code can be derived directly from the model code, which has two main advantages. It reduces the complexity of the construction of the adjoint model and it avoids the inconsistency that can arise from the derivation of the adjoint model followed by its discretization due to non-commutativity of adjoint and discretization operations. Navon *et al.* (1992a) and Talagrand (1991) derived the adjoint model code from the tangent linear model code while Lawson *et al.* (1995a) used the Lagrange multiplier approach to derive the adjoint code from the direct model code. The variational adjoint assimilation algorithm includes four parts: the direct model, the construction of the adjoint code of the forward model, the computation of the cost function and its gradient with respect to the control parameters, and the large-scale unconstrained local optimization algorithm.

The objective of this study is to assimilate tide gauge observations in order to estimate the important model parameters, *i.e.*, bottom and wind stress, and get the best representation of the circulation in the Bay. This goal was achieved by using a two-dimensional (2-D), vertically-integrated shallow water equations model. Several studies for the English Channel (Ozer and Jamart, 1988; Jamart and Ozer, 1989; Werner and Lynch, 1987) have indeed recognized that the 2-D, inviscid shallow water equations can provide an accurate solution to the problem of tidal propagation. The assimilation technique is the variational adjoint method where the adjoint model code is obtained from the tangent linear version of the model code. The minimization algorithm used to minimize the cost function is the limited memory quasi-Newton method developed by Gilbert and Lemaréchal (1989) which is similar to the method developed by Liu and Nocedal (1989).

Specifically, the results obtained from the assimilation study address the following questions:

- How can variational data assimilation be used to determine the forcing in the model, *i.e* wind stress and bottom friction? Can we estimate the spatial and/or time dependence of the bottom friction and wind stress using tide gauge data?

- Is the number of tide gauges adequate to predict the sea level in the bay? Are the gauges well distributed around the bay?

- Can variational data assimilation be used to determine the adequacy of a two-dimensional model to reproduce the main features of the circulation in the Chesapeake Bay?

Section 2 contains background information on the physics and the circulation of

the Chesapeake Bay. Section 3 contains details of the variational adjoint data assimilation method, the direct model and application of data assimilation to the Bay. The results of surface elevation assimilation in the case of identical twin experiments, tidal and wind-driven circulation experiments using the tide gauge observations are given in Section 4 while they are discussed in Section 5. Conclusions of this study are presented in Section 6.

2 Background

The first question that might be asked is "what is an estuary?". Historically, the term estuary comes from the latin name *aetus*, which means tide, and applied to the lower tidal reaches of rivers. Cameron and Pritchard (1963) extended the definition as follows: " An estuary is a semi-enclosed body of water which has a free connection with the open ocean and within which seawater is measurably diluted with fresh water derived from land drainage". Estuaries have been classified based on two criteria: 1) geomorphological properties (Pritchard, 1952a; Dyer, 1973) and 2) circulation and stratification patterns (Dyer, 1973). In terms of shape, an estuary can be of three types: coastal plain, deep basin, and bar-built estuary. In terms of the water properties, an estuary can be classified as highly stratified salt wedge type, highly stratified, partially mixed and vertically homogeneous. Complete descriptions of estuary types can be found in Dyer (1973) and Pickard and Emery (1982). In the past decades, the main focus of estuarine studies has been on tidal and gravitational circulation as well as river runoff. Only recently, it has been acknowledged that wind-driven circulation might at times be more important than the gravitational circulation. In the following sections, the physical characteristics of the Chesapeake Bay, the field observations, and the tidal and wind-driven circulation in the Bay are briefly described.

2.1 Physical characteristics of the Chesapeake Bay and its tributaries

Chesapeake Bay is the largest estuary in the United States with a length of roughly 310 km, a width averaging 30 km (Fig. 1) and a complex topography. The Bay can be divided into two regions which have different dimensional and physical characteristics: the main stem and the tributaries (over 50). The main stem which is very narrow at the entrance (18.5 km) widens to about 35 km near the mouth of the Potomac River. It then narrows to about 6 km near the mouth of the Severn River. Its main axis is directed north-south except in the lower part where it is in the northwest-southeast direction. The main stem has an average depth of 8 m with a maximum depth, nearly 53 m, off Kent Island. The 18 ft (5.5 m) and 36 ft (11 m) depth contours are shown in Fig. 1. The depth contours show that there is a complicated system of channels starting at the Bay mouth and branching into the tributaries. Two main channels and a third narrow channel start at the entrance of the Bay. The southern channel in the entrance extends westward as Thimble Shoal channel and branches up the James River. The main channel extends into the main stem while the northern channel, very narrow and deep, extends northward along the eastern shore into the Pocomoke sound. The major tributaries of the lower Bay are the Rappahannock, York, and James Rivers which account for approximately 20% of the freshwater input in the Bay. The Potomac and Susquehanna Rivers are the major tributaries of the upper Bay, with the Susquehanna accounting for about 50% of the total freshwater input.

Based on geomorphological properties (Pritchard, 1952a), the Chesapeake Bay has been classified as drowned river valley or coastal plain estuary. After the glacial period, roughly 10 thousand years ago, the Chesapeake Bay system was formed. Before sea level rose about 100 m following the glacial period, the Susquehanna River reached the ocean about 180 km seaward of the present shoreline and the York and

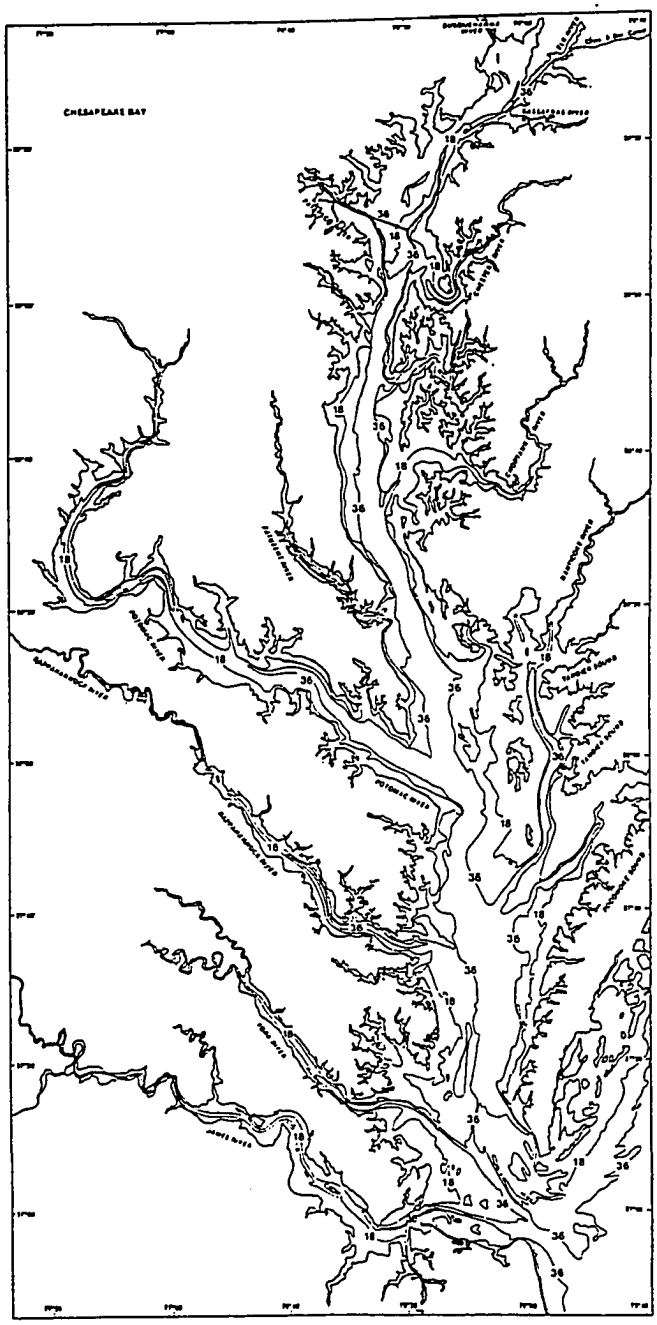

Figure 1: Depth contours of the Chesapeake Bay expressed in feet (Fisher, 1986).

other rivers were tributaries of the Susquehanna River. Based on the circulation and stratification patterns, the lower Bay has been classified as vertically homogeneous with lateral variations of salinity during normal runoff conditions (Pritchard, 1952b). In that region, tidal flow is dominant. The upper Bay and the tributaries have been classified as slightly stratified. Due to the river discharge and the tide, a two layer circulation is present with a net seaward flow of fresh water in the upper layer and a net flow of saline water toward the head in the lower layer.

2.2 Observations in the Chesapeake Bay

Observations are required not only to understand the circulation in the Bay but also to verify the success of the numerical models and to estimate model parameters using data assimilation techniques. The most useful information comes from synoptic data. The Chesapeake Bay has now been monitored for decades by several laboratories and universities and measurements include sea level, current, temperature, salinity, water quality as well as meteorological observations. Below is a brief review of the publically available observations.

- *Water level and current*

For more than a century, tide and tidal currents have been observed in the Chesapeake Bay. The first tide station was installed in Annapolis in 1844 (Haight *et al.*, 1930; Hicks, 1964; Fisher, 1986). Prior to 1964, more than 200 tide gauge stations and over 100 near-surface current stations were deployed. However, they were not usually deployed for a long time or at the same time. From those stations, ten tide gauge stations in the Chesapeake Bay and its tributaries are now part of the National Tide and Water Level Observation Network (Table 1) and are permanent installations maintained by the National Oceanic and Atmospheric Administration (NOAA). In addition to those long term measurements, two extensive tide and cur-

rent surveys of the Chesapeake Bay were conducted from 1970 to 1974 and from 1981 to 1983 (Fig. 2) by the National Ocean Survey to update tide and tidal current predictions and to provide tidal datum for shoreline boundary determination.

• *Meteorological observations*

For a long time, meteorological observations were collected only at the major airports, *e.g.*, Baltimore, Washington DC, Norfolk International Airports and Patuxent River Naval Air Station. It is only recently that meteorological observations became available over the water. Starting in 1985, two buoys were deployed in the Chesapeake Bay by the National Data Buoy Center (NDBC) as part of the Coastal-Marine Automated Network (C-MAN) program. The first buoy is located in the upper Bay at Thomas Point, Maryland (38.9° N, 76.4° W) while the second one is located outside the Bay at the Chesapeake Light Tower, Virginia (36.9° N, 75.7° W). Wind speed, direction and gust, barometric pressure and air temperature are processed every hour and transmitted to the users. In addition to those buoys, meteorological observations are available at some tide gauge stations, *e.g.*, at the Chesapeake Bay Bridge Tunnel (CBBT).

• *Water properties*

From 1985 to 1991, water quality data were collected at more than 130 stations in the main stem and the tributaries. Monitoring in the main stem was part of a joint program between University of Maryland, Old Dominion University, and Virginia Institute of Marine Science and was supported by the U.S. Environmental Protection Agency (EPA). Monitoring in the tributaries was done by state regulatory agencies. This comprehensive data set is now available on CD-ROM (Rennie and Neilson, 1994). In addition, sea temperature is routinely measured at the aforementioned buoys and some of the tide gauge stations.

National Ocean Service long term control tide stations

Station Number	Station Name	Latitude (N)	Longitude (W)	Installation date
8574070	Havre de Grace, MD	39°46.9′	76°05.5′	1971
8574680	Baltimore, MD	39°16.0′	76°34.7′	1902
8575512	Annapolis, MD	38°59.0′	76°28.8′	1929
8571890	Cambridge, MD	38°34.5′	76°04.3′	1942
8577330	Solomons Is, MD	38°19.0′	76°27.2′	1938
8635750	Lewisetta, MD	37°59.8′	76°27.8′	1970
8637624	Gloucester Pt, VA	37°14.8′	76°30.0′	1950
8632200	Kiptopeake, VA	37°10.0′	75°59.3′	1951
8638610	Hampton Roads, VA	36°56.8′	76°19.8′	1927
8638863	CBBT, VA	36°58.1′	76°06.8′	1975

Table 1: Tide gauge stations part of the National Tide and Water Level Observation Network.

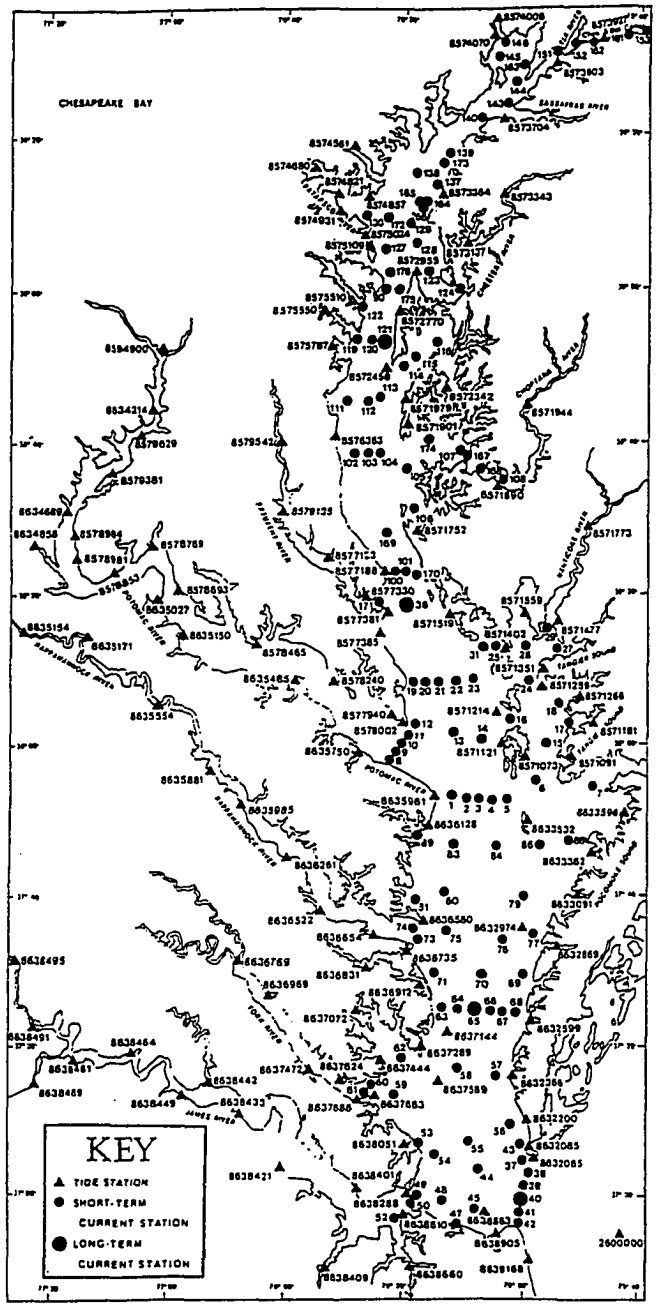

Figure 2: Overview of tide and current station deployments in the Chesapeake Bay and its tributaries (Fisher, 1986; Browne and Fisher, 1988).

• *Bathymetry*

Bathymetry of the Chesapeake Bay is available from the National Ocean Service (NOS) hydrographic data base. The data set includes depths for the main stem (except a small portion north of Baltimore) as well as the major tributaries. The bathymetry is available on a 15-second grid.

2.3 Circulation in the Chesapeake Bay

Estuarine circulation, due to the combined effects of tide action, horizontal salinity gradients, river runoff and meteorological forcing (wind stress, inverted barometer effect), has been intensively studied in the Chesapeake Bay and its tributaries and is still an ongoing source of research activity. Diverse investigations from field observations, *e.g.*, temperature, salinity and current, and from simple models have been carried out by Pritchard and other researchers in order to explain the gravitational and tidal circulation. A summary of those studies can be found in Officer (1976). It is only during the last two decades that wind-driven circulation has been shown to be as important as the gravitational circulation, indeed the dominant non-tidal circulation at times. For example, Weisberg (1976) found that in the Providence River of the Narragansett Bay, wind effects can be of equal or greater importance to the tidal or gravitational circulation. Since our study focuses on the barotropic circulation, only tidal and wind-driven circulation will be discussed in the following sections.

2.3.1 Tidal circulation

Description of the tidal circulation of the Chesapeake Bay from sea level and current measurements started with Harris (1907) and was further investigated by Hicks (1964). They were able to construct approximate cotidal and co-current charts for the main stem and the tributaries. Their study showed that the dominant tidal

constituents in the Bay are the semidiurnal, M_2, and the diurnal, K_1, constituents. Hicks (1964) also found that the Chesapeake Bay can contain one complete wavelength of a semidiurnal tidal wave and a half wavelength of a diurnal wave. More recently, an extensive analysis by Fisher (1986) of the sea level and current data collected by NOAA at 108 and 124 locations (Fig. 2), respectively, during two surveys, from 1970 to 1974 and from 1981 to 1983, gave more insight into the details of the tidal circulation. Charts of the cophase and coamplitude lines for M_2 and K_1 tides are shown in Figs. 3 and 4. One degree in phase of an M_2 tidal cycle corresponds approximately to two minutes in time while one degree in phase of an K_1 tidal cycle corresponds to four minutes. Charts of the cospeed and cophase of the M_2 tidal current are plotted in Figs. 5 and 6.

Several main features of the tidal circulation can be pointed out from these charts. First, the coamplitude line configuration reflects the expected effect of the earth's rotation which is manifested by a larger amplitude of the tide on the Eastern Shore than on the Western Shore. The M_2 and K_1 tides are Kelvin waves, which is suggested by the pattern of orthogonally-oriented cophase and coamplitude lines in the lower Bay. The nature of the waves is further supported by the location of the minimum of the M_2 amplitude near the Potomac River, three-quarters of an M_2 wavelength from the head of the Bay, which is consistent with the pattern expected from the superposition of incident and reflected Kelvin waves damped by friction. The reflected wave can also be seen in the rapid decrease of the current north of Havre de Grace. An increasing effect from the north end of the Bay is found in the phase difference between tidal elevation and current. At the entrance of the Bay, the tidal elevation leads the tidal current while in the middle of the Bay they are in phase. At the head of the Bay, the tidal current leads the tidal elevation.

The effect of the bottom friction and topography is further seen in the configuration of the coamplitude and cophase lines of the M_2 and K_1 as well as in the pattern

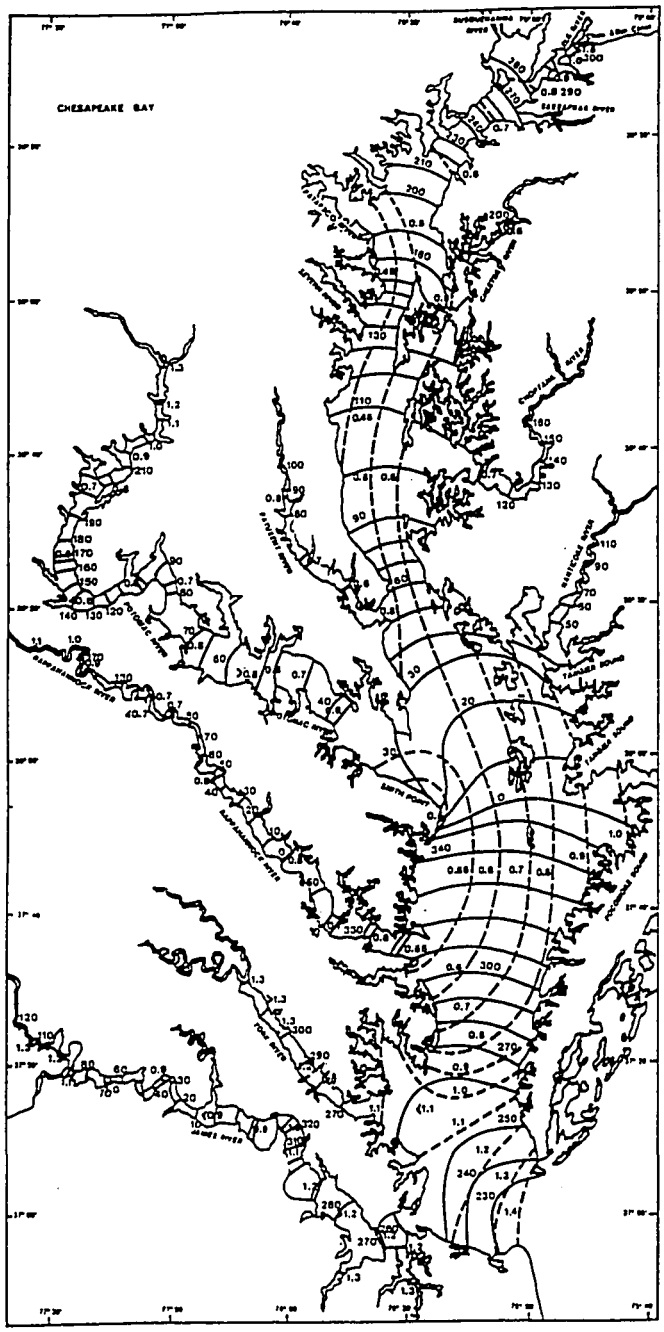

Figure 3: Superposition of coamplitude and cophase lines of M_2 tide (Fisher, 1986). Cophase lines (solid) are expressed in degrees and coamplitude lines (dashed) in feet.

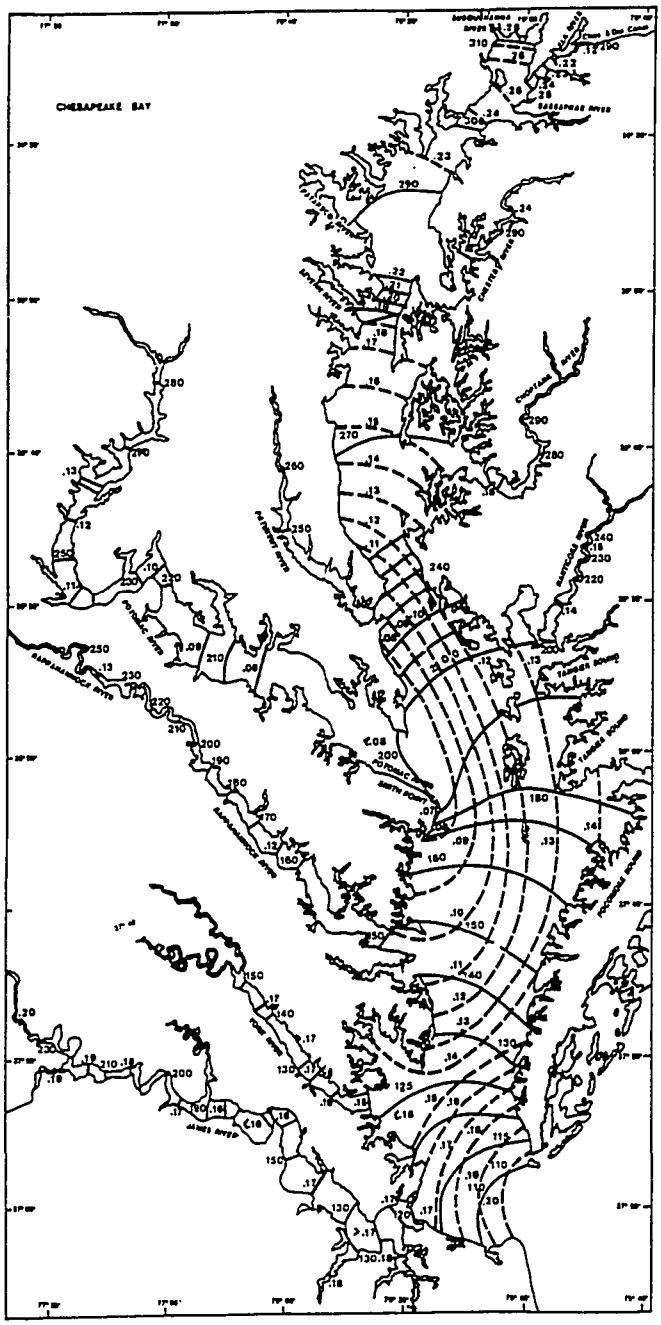

Figure 4: Superposition of coamplitude and cophase of K_1 tide (Fisher, 1986). Cophase lines (solid) are expressed in degrees and coamplitude lines (dashed) in feet.

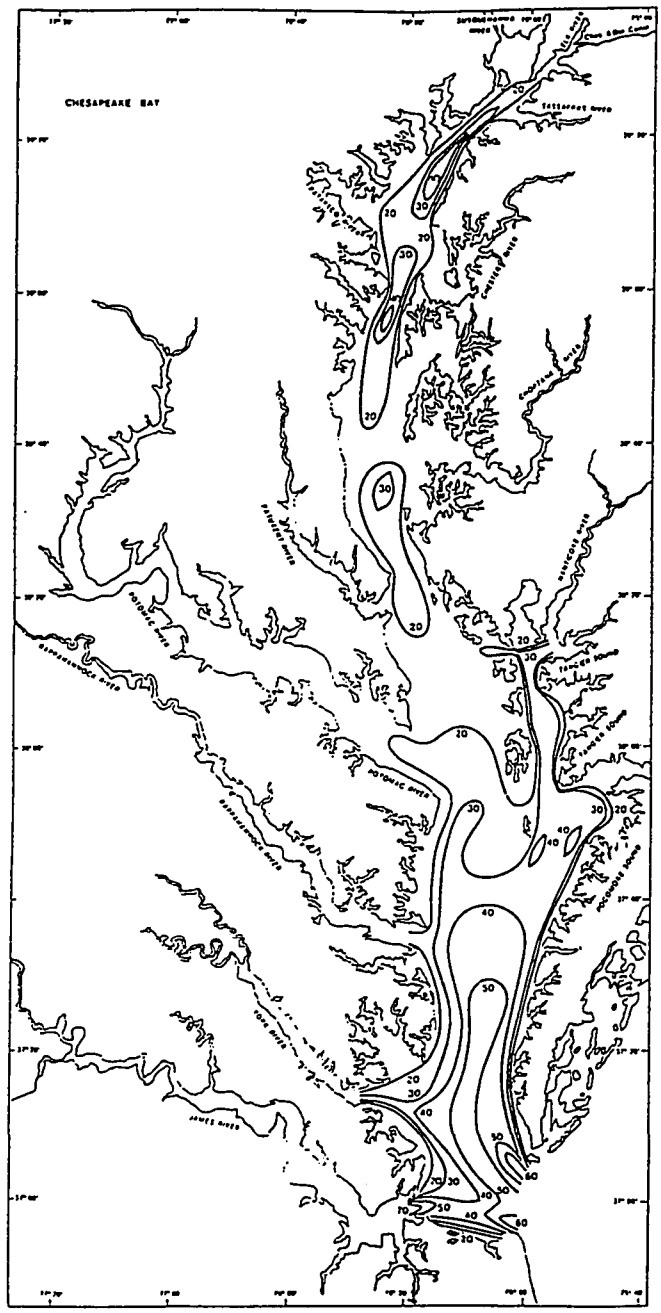

Figure 5: M_2 tidal current cospeed lines expressed in centimeters per second (Fisher, 1986).

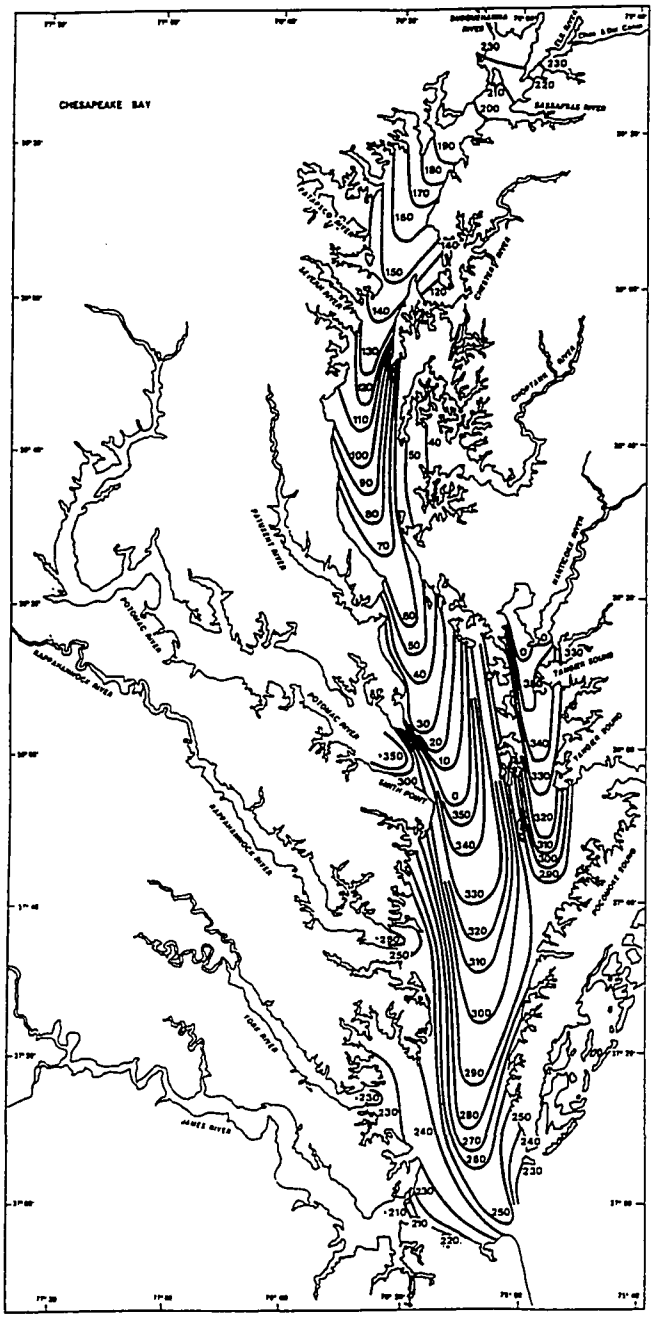

Figure 6: M_2 tidal current cophase lines expressed in degrees (Fisher, 1986).

of the cophase of the tidal current. Two virtual amphidromic points (Defant, 1961), characterized by the convergence of the cophase lines of the M_2 and the concentric nature of the coamplitude lines with an outwards increasing amplitude are found near the Potomac River (north of Smith Point) and west of the Severn River. The migration of the amphidromic point towards the Western Shore and the Bay's entrance is the result of the bottom friction (Fisher, 1986). An amphidromic point north of Smith Point is also found for the K_1 tide. While the M_2 tidal current cophase pattern differs significantly from that of the M_2 tide throughout the main stem, the curvature of the tidal current cophase lines towards the entrance of the Bay is mainly due to bottom friction. The effect of the topography can be pointed out by the presence of two flood currents at the entrance of the Bay which are located in the two channels. In the southern channel, the current splits in two parts, *i.e.*, the major current proceeds northward up the Bay and the smaller one is directed towards the James River. In the northern channel, the speed decreases very quickly up the Bay. North of the Patuxent River, the cospeed lines are closed which also suggests a strong effect of the bottom topography.

The M_2 and K_1 cophase and coamplitude lines in the major tributaries are mainly oriented cross channel. The phase in the lower region of the rivers increases rapidly while it changes very little with distance in the upper part near the limit of tide. This suggests that the wave is more like a progressive wave in the lower part and a standing wave in the upper region. Fisher (1986) also shows that there is little amplification of the elementary tidal constituents, *e.g.*, M_2 and K_1. However, the shallow water constituents, which include the overtides and the compound tides, are significantly amplified near the limit of tide in the major tributaries.

As shown in the study from Fisher (1986), the bottom friction is one of the major forcings for the tidal circulation. Yet, this forcing is little-known and hard to measure. As shown in the next sections, by assimilating surface elevations from tide

gauges, it is possible to get an estimate of the bottom drag coefficient and therefore the bottom stress.

2.3.2 Wind-driven circulation

Non-tidal circulation in partially mixed estuaries can be driven by a horizontal density gradient, wind forcing and river runoff. While the gravitational circulation was thought during the past decades to be the main component of the non-tidal circulation and was extensively studied in the Chesapeake Bay, it has been shown that the wind-driven circulation can at times be larger than the gravitational circulation. Wind-driven circulation has mainly been studied from field observations.

While the response of the water to the wind forcing in the Bay is complex, several studies have shown that sea level fluctuations depend on local winds (local forcing) and exchange between coastal ocean and estuary (non-local forcing). Sea level measurements for a period of two months in 1974, show non-tidal fluctuations at period of 20, 5 and 2.5 days (Fig. 7) (Wang and Elliott, 1978). The 20-day sea level fluctuations, with amplitude decreasing towards the head of the Bay and phase increasing up Bay, were found to be the result of up-Bay propagation of coastal sea level fluctuations generated by alongshore winds. Fluctuations of 5-day period, with amplitude almost uniform in the Bay, were driven by coastal sea level fluctuations and local cross-shore winds. Seiche oscillations at 2.5 day periods, with amplitude larger at the head of the Bay and with constant phase within the Bay, were driven by the local longitudinal winds. Wang (1979a) extended the previous study to one year of sea level measurements and found barotropic responses at similar periods. Contrary to the study by Wang and Elliott (1978), Wang (1979a) found that the driving force at 10-day period was not the coastal alongshore forcing but instead the cross-shore wind. He also observed that the seiche oscillations were intensified in winter due to the passage of extratropical cyclones. More recently, Valle-Levinson

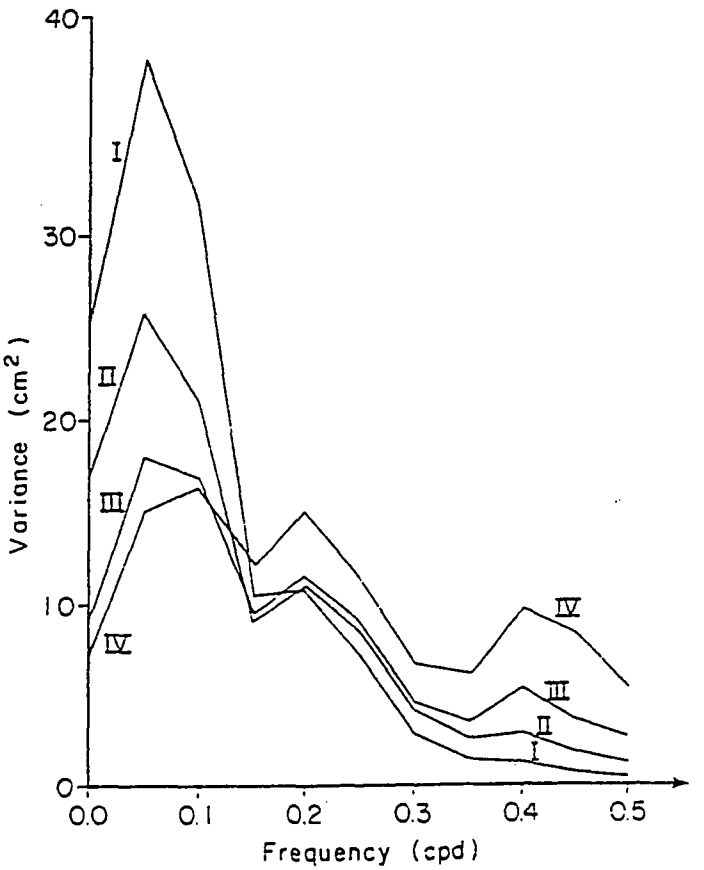

Figure 7: Power spectra of sea level, at (I) Kiptopeake Beach, (II) Lewisetta, (III) Solomons Island and (IV) Annapolis. (Wang and Elliott, 1978).

(1995) and Paraso and Valle-Levinson (1995) studied the atmospheric response on the barotropic exchange in the lower Bay. From observations of sea level, wind and sea temperature, they showed that northeasterly winds cause net barotropic inflows at the Bay entrance while southwesterly winds cause net outflows. A more rapid change in the sea level was also noticed when the wind blows from the north-east.

The influence of the wind on the vertical structure of the nontidal circulation has also been investigated. A study by Pritchard and Rives (1979) in the middle reaches of the Chesapeake Bay showed that the subtidal nongravitational currents were wind-driven in the upper layer of the column of water while in the deeper layer the currents were in the opposite direction. Based on current measurements during winter 1975, Wang (1979b) reported a strong wind-driven barotropic circulation in the lower Bay and a baroclinic circulation in the upper Bay which he related to two forcing mechanisms, *i.e.*, surface slope of the water and the wind. He further explained this flow pattern using a conceptual frictional model. The effect of the wind on the vertical structure of the residual currents in the middle reaches of the Bay was further investigated by Vieira (1986). Based on current, tidal elevation and wind measurements, Vieira (1986) concluded that the upper layers (8 m) are directly driven by the wind while in the lower layers the flow is in the opposite direction of the wind as a result of the surface slope associated with it. Finally, several events leading to the destratification of large areas of the Bay has been studied by Goodrich *et al.* (1987) from current and salinity observations between 1981 through fall 1983. They found that those events mainly occur from early fall through mid-spring. Blumberg and Goodrich (1990) further investigated a specific event during September 1983 using a three-dimensional numerical model.

While general patterns of the wind-driven circulation have mostly been studied, isolated events of wind-driven circulation have also been pointed out by Chuang and Boicourt (1989). An abnormal seiche motion was detected in April 1986. The

seiche lasted for seven days and had a period of 1.7 days which is shorter than the natural period of the Bay. The seiche was found to be generated by a cross-shore wind at the mouth of the Bay. The wind-driven current was then able to initiate a resonant seiche in the rest of the Bay. Those conclusions were also supported by a simple analytical model. Chao (personal communication) developed an analytical model for a L-shaped estuary such as the Chesapeake Bay in order to explain the April 1986 event. He showed that for L-shaped estuaries, both locally and remotely forces responses are important.

Based on the previous studies, it appears that the wind-driven circulation is important in the Chesapeake Bay and is very complex. However, the wind data are mainly available at the major airports and conversion of wind on land to wind over water is not an easy task. Chao (personal communication) showed that in order to reproduce the event found in April 1986 the longitudinal wind used in the Bay has to be increased compared to the wind measured at Norfolk airport. This increase of wind over water compared to that measured over land has also been pointed out by Wong and Garvine (1984), Goodrich (1985). To explain the sea level in the Delaware Bay, Wong and Garvine (1984) had to increase the shore-based wind stress fourfold. Goodrich (1985) found that while the longitudinal winds attenuate rapidly toward the shores, the lateral winds do not, and that over-water/over-land regression slopes for north and east components of the wind are 2.5 and 1.43, respectively. In the next sections, we shall show how it is possible to estimate the wind forcing in the Bay from assimilation of tide gauge observations and get the best representation of the wind-driven circulation.

3 Method

3.1 Overview

During the past decade, interest in data assimilation has increased in meteorology as well as in oceanography. One of the reasons for developing data assimilation methods in meteorology was the need for weather prediction. Meteorologists are now routinely assimilating atmospheric data in numerical weather forecasting systems. While the need of data assimilation in oceanography is generally not for prediction but rather for understanding the ocean, data assimilation has become an important tool for the oceanographers with the availability of larger data sets, development of new observational techniques (*e.g.*, altimeters, satellites, tomography), and increases in computing power. Reviews of data assimilation can be found in Ghil *et al.* (1981), Bengtsson *et al.* (1981), Lorenc (1986), Navon (1986), Haidvogel and Robinson (1989), and Ghil and Malanotte-Rizzoli (1991).

Data assimilation techniques can be divided into two main classes: statistical interpolation and variational analysis. Statistical interpolation methods which include successive correction (Cressman, 1959; Bratseth, 1986) and optimal interpolation (Gandin, 1963; Lorenc, 1981) are routinely applied in weather forecasting. In optimal interpolation, the model results are corrected by adding a weighted fraction of the difference between model results and data. The weight is determined from the error covariance of the model. An extension of the optimal interpolation method is the Kalman or Kalman-Bucy filtering (Kalman, 1960; Kalman and Bucy, 1961), where the correction to the model solution is obtained by computing the error co-

variance function, *i.e.*, the covariance of the differences between model results and data, from the numerical model dynamics. This method is more accurate than the classical optimal interpolation but it requires considerable computing power. Several applications of Kalman filtering can be found in oceanography (Miller, 1986; Budgell, 1986, 1987; Bennett and Budgell, 1987; Heemink and Kloosterhuis, 1990; Fukumori *et al.*, 1993). The main weakness of these methods is that, for non-linear models, it is difficult to estimate the model and observation error covariance matrices.

The second class of data assimilation methods, the variational analysis, has as its objective to find the assimilating model solution which minimizes a predefined objective function. This function, called the cost function, measures the distance between model solutions and observations. The main difference between variational technique and interpolation methods is that the analyzed field must satisfy an explicit dynamical constraint which is usually expressed by the model equations. Furthermore, variational assimilation has two main advantages: it can be applied to linear as well as non-linear models; and, as shown in the next sections, it can be implemented in a straightforward manner.

Calculus of variations was introduced in meteorology by Sasaki (1955, 1970) who introduced the concept of "weak" and "strong" constraints, *i.e.*, conditions imposed on the resulting flow field. Since then, this method has been considerably developed in both meteorology and oceanography. For example, Bennett and McIntosh (1982) and Bennett (1985) used variational methods to investigate open boundary conditions in tidal model and array design. Schröter and Wunsch (1986) studied the effect of data errors in oceanic circulation models. In the same time, methods to compute the gradient of the cost function, which is required during the minimization process, were going through considerable development. The adjoint method is the most powerful tool to compute the gradient of the cost function with respect to the control

variables. The variational adjoint technique then leads to the question of sensitivity analysis, variational adjustment and parameter estimation. For example, Cacuci (1981), Hall *et al.* (1982), Hall and Cacuci (1983) and Cacuci and Hall (1984) used adjoint models to assess sensitivity of model forecasts to changes in boundary conditions and model parameters. Lacarra and Talagrand (1988) identified regions of enhanced barotropic and baroclinic instability while Farrell and Moore (1992) tried to find the fastest growing unstable modes of an oceanic jet. Zou *et al.* (1993b) examined the sensitivity of a blocking index in a 2-layer primitive equation isentropic spectral model.

Variational adjustment found its first applications in meteorology for weather forecasting. While first limited to barotropic models (Courtier, 1984; Lewis and Derber, 1985; Courtier and Talagrand, 1987; Talagrand and Courtier, 1987), the method has been extended in the recent years to more complicated models (Thépaut and Courtier, 1991; Navon *et al.*, 1992; Thépaut *et al.*, 1993; Courtier *et al.*, 1993). In oceanography, variational adjustment was first applied with simple models. For example, Long and Thacker (1989a,b) used a simple equatorial-wave model to recover the model state from surface elevation and wind stress observations, while Sheinbaum and Anderson (1990) used a one-layer, linear, reduced-gravity model. With the increase of computer power and of data availability (*e.g.*, from Geosat altimeter and satellites), variational adjustment became extensively used. Moore (1991) assimilated data into a quasi-geostrophic, multi-layer, open-ocean model of the Gulf Stream region and estimated initial conditions. Tziperman *et al.* (1992a,b) used a general circulation model and estimated values of the model inputs consistent with a steady circulation and available data in the North Atlantic Ocean. Seiler (1993) used a similar quasi-geostrophic model for estimation of the open boundary conditions. Greiner and Périgaud (1994) assimilated Geosat sea-level variations into a nonlinear shallow-water model of the Indian Ocean. These studies only represent

a limited sample of what has been done in the field of variational adjustment.

Parameter estimation also became a field of interest for variational adjoint methods. Panchang and O'Brien (1989) used the adjoint method to determine the bottom friction coefficient in a channel. Smedstad and O'Brien (1991) estimated the effective phase speed in a model of the equatorial Pacific ocean using sea level observations, while Yu and O'Brien (1991) used the same technique to estimate the eddy viscosity and surface drag coefficient from an observed velocity field. This work has been extended by Richardson and Panchang (1992), and Lardner and Das (1994). Das and Lardner (1991, 1992) and Lardner (1993) used a two-dimensional tidal model and a variational method to estimate bottom drag, depth and open boundary conditions. Similarly, Lardner *et al.* (1993) estimated the bottom drag coefficient and bathymetry correction for a two-dimensional tidal model of the Arabian Gulf.

Finally, a new field of application of the adjoint variational method, marine system modeling, has captured interest. Lawson *et al.* (1995a) applied the variational adjoint method to a prey-predator model to assess the recovery of little-known biological parameters, such as initial concentrations, growth and mortality rates. It was shown that the ease of recovering initial concentrations and rates depends not only on data availability but also on the form of the model equations. Based on identical twin experiments, the recovery of initial concentrations and rates was possible even with a data set containing only information on either prey or predator abundance. However, when only the abundance of one species was available, the structure of the biological model, *i.e.*, the process that couple ecosystem components, had to be modified. In a second study, Lawson *et al.* (1995b) applied the adjoint method to a five-component time-dependent ecosystem model and recovered population growth and death rates, amplitudes of forcing events and component initial conditions. The effect of data distribution and data type on the ability to recover model parameters was investigated using identical twin experiments and sampling strategies corre-

sponding to US. JGOFS experiments at the Bermuda Atlantic Time Series (BATS) and the Hawaii Ocean Time Series (HOT) stations.

The approach chosen to assimilate tide gauge data in the Chesapeake Bay and recover wind, bottom forcing and circulation in the Bay is the variational formalism, often referred to as the adjoint method. As mentioned earlier, variational data assimilation consists in finding the model solution which minimizes an objective function, the cost function, measuring the difference between the model solution and the available data. Various optimization procedures can be used to determine the minimum of the cost function. Most of the optimization techniques require the computation of the gradient of the cost function with respect to the control variables, which can be achieved by direct perturbation (Hoffman, 1986) but the numerical computation is very costly. A powerful mathematical tool used to numerically compute the gradient of the cost function is the adjoint of the equations of the assimilating model. As shown in the coming subsections, the derivation of the adjoint equations can be achieved by various methods.

The variational adjoint method then includes four components: the mathematical model (circulation model) or forward model, the computation of the cost function from the data and model output, the adjoint of the forward model or backward model and an optimization technique. The four components are used in an iterative process which leads to the determination of the control variables giving the best fit to the data (Fig. 8) and can be described as follows. The direct model is run using a first guess of the control variables. The model output and data are then used to compute the value of the cost function. Thereafter, the adjoint of the model is used to compute the gradient of the cost function with respect to the control variables, which is then used in the large scale unconstraint local optimization procedure to compute the search direction towards the minimum and the optimal step size in that direction. A new value of the control variables is then estimated and the model

is rerun. This procedure is applied until a preset convergence criterion is satisfied, *e.g.*, $\|\nabla J\| \leq \epsilon_1$ and/or $J \leq \epsilon_2$ where ϵ denotes a small value and J is the cost function. The four steps of the assimilation technique are described in the following sections.

3.2 Circulation model

The circulation model used to study the barotropic circulation in the Chesapeake Bay is a conventional 2-D vertically-integrated shallow water equations model. This model was developed by MUMM (Management Unit of the Mathematical Models of the North Sea and Scheldt Estuary) to study the tidal propagation in the English channel (Ozer and Jamart, 1988; Jamart and Ozer, 1989) and will be referred to as the MU-model.

In a right handed coordinate system, with the z-axis pointing upwards, the governing equations are :

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} - fv = -g\frac{\partial \eta}{\partial x} + \frac{\tau_w^x}{\rho H} - \frac{\tau_b^x}{\rho H} \tag{1}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + fu = -g\frac{\partial \eta}{\partial y} + \frac{\tau_w^y}{\rho H} - \frac{\tau_b^y}{\rho H} \tag{2}$$

$$\frac{\partial \eta}{\partial t} + \frac{\partial (Hu)}{\partial x} + \frac{\partial (Hv)}{\partial y} = 0 \tag{3}$$

where t denotes time, f the Coriolis parameter, g the acceleration due to gravity, τ_w^x and τ_w^y the components of the wind stress, and H the total water depth. The unknown η is the elevation of the free surface with respect to the mean sea level, and u and v are the components of the vertically-averaged velocity. The bottom stress $\vec{\tau_b}$ is parameterized by means of a quadratic dependence with respect to the depth mean current,

$$\vec{\tau_b} = \rho c_D \|\vec{u}\| \vec{u}. \tag{4}$$

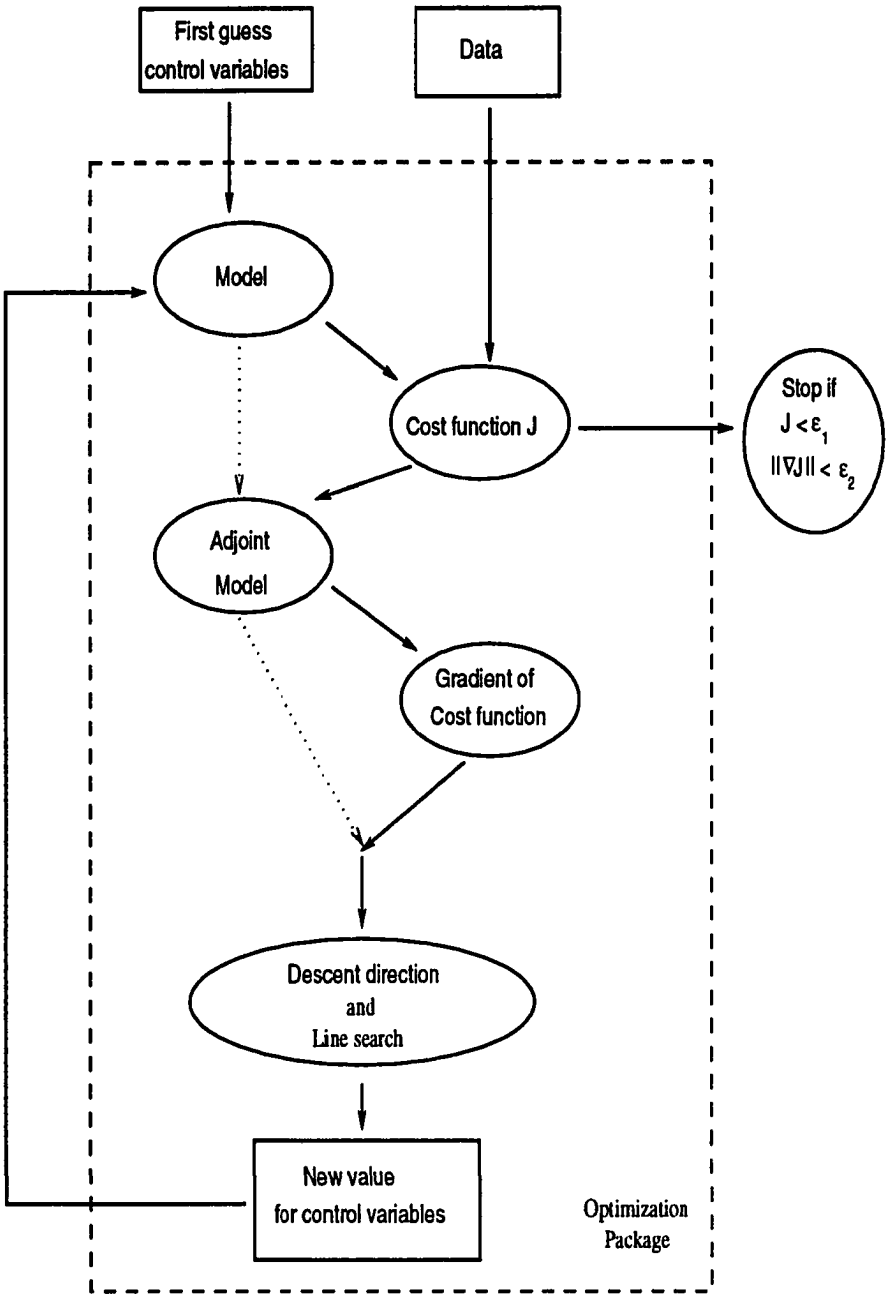

Figure 8: Schematic of the steps involved in the data assimilation scheme. The solid lines indicate the main path taken during the procedure.

In practice, the bottom drag coefficient c_D varies with water depth, seabed composition and phase of the tide. It is parameterized as

$$c_D = \frac{g}{c^2} \text{ with } c = \frac{H^\alpha}{n} \tag{5}$$

where c ($m^{1/2}s^{-1}$) is the Chezy coefficient and n is the Manning's roughness (Officer, 1976). Typical values for α and n are respectively 1/6 and 0.02 giving a drag coefficient of $\sim$ 0.002 for a depth of 10 m. Other parameterizations of the bottom drag coefficient will be described in Section 5.

Along the open boundaries which coincide with a grid line containing elevation points (see grid definition, Fig. 9), the free surface elevation is imposed. For instance, the forcing due to the tide is introduced by specifying the time evolution of the free surface as follows:

$$\eta(\vec{s},t) = \sum_{k=1}^{K} f_k a_k cos(\omega_k t + (V_o + u)_k - \phi_k) \tag{6}$$

where $\vec{s}$ denotes the position of a point along the boundary, a_k and ϕ_k are the harmonic constants (amplitude and phase) of the kth constituent, ω_k is the frequency of the kth constituent, f_k is a factor to reduce the mean amplitude of the constituent to the starting time of the simulation, $(V_o + u)_k$ is the value of the equilibrium argument of the kth constituent at the starting time of the simulation, t denotes the time elapsed from the beginning of the computational time. Due to the presence of the advection terms in the momentum equation, an additional boundary condition is necessary at those times when the water flows towards the interior of the domain. In that case, the gradient of the depth mean current in the direction perpendicular to the boundary is set equal to zero along the boundary grid line and the grid line half a Δx or a Δy inside the domain.

Along the solid boundaries which coincide with a grid line containing velocity unknowns, the component of the total transport in the direction perpendicular ($\vec{n}$) to the boundary is set equal to zero. The additional condition ($\frac{\partial \vec{u}}{\partial n} = 0$) is applied

along a grid line a half Δx or a Δy inside when the water flows towards the interior of the domain.

The equations are solved by means of finite difference analogs (Ozer *et al.*, 1990) on a uniform staggered grid (Arakawa C-grid) (Fig. 9) . The time stepping scheme is a semi-implicit, alternate direction method (ADI) (Beckers and Neves, 1985), which is unconditionally stable and allows larger time steps than permissible by other explicit time differencing schemes such as leapfrog. The ADI scheme can be summarized as follows.

During the first half time-step, the north-south (y) component of the momentum equation is solved explicitly, except for the bottom stress term. Thereafter, the east-west (x) component of the momentum equation is solved by taking the surface gradient term at the new time level. The continuity equation is then solved with the x-component of the depth mean current at the new time level. Due to the linearity of the finite difference approximation of the x-component of the momentum during this step, the x-component of the velocity is decomposed into two components. The first component is the solution of the momentum equation without the surface gradient term and is already obtained. The second component is proportional to the unknown surface gradient term. Its finite difference approximation is then introduced into the finite difference analog of the continuity equation which leads to a tridiagonal algebraic system for the surface elevation at the new time level. Once the value of the free surface is known at the new time level, the x-component of the velocity can be computed. The procedure during the second half-time step is entirely similar except that the implicit procedure is applied in the y-direction.

3.3 Cost function

The variational adjoint method attempts to find a set of model control variables, *e.g.*, initial conditions at the start of the assimilation window, open boundary con-

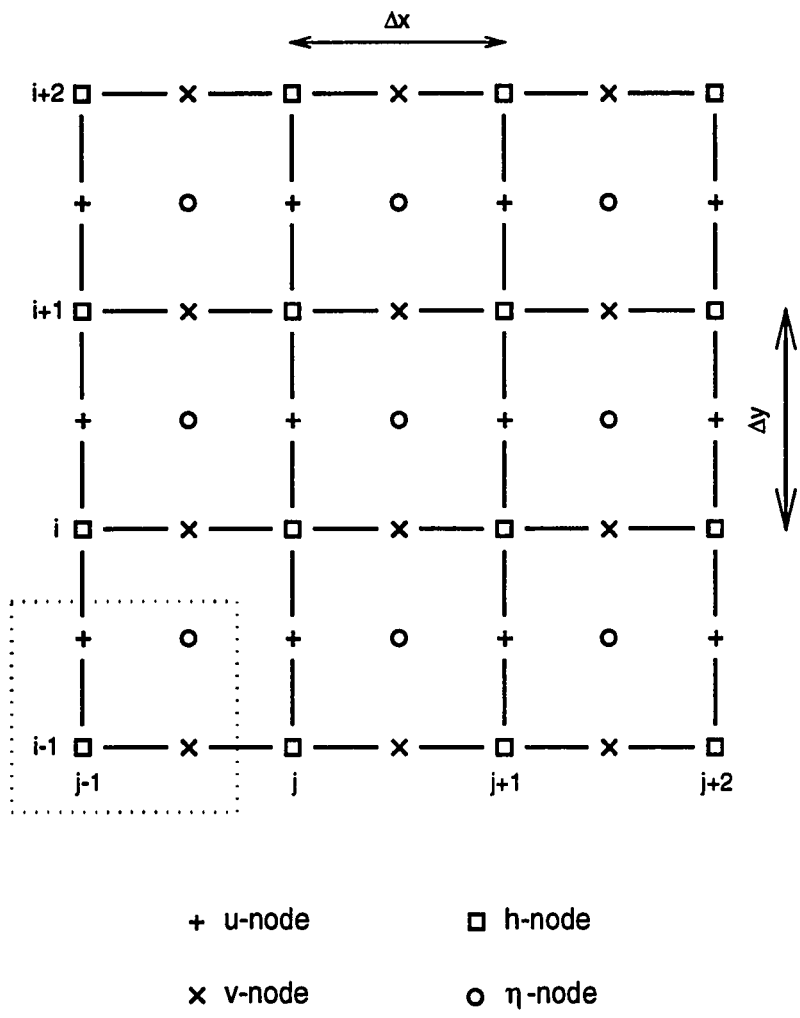

Figure 9: Staggered grid and indexing used in the model. The variables inside of the dotted box have the same value of the indexes i and j.

ditions, parameters of the model, most consistent in a least-squares sense with the observations over a given period of time, *i.e.*, the window of assimilation. This is achieved by computing the minimum of a cost function with respect to the control variables. Thus, the first task is to define a cost function suitable for the study under consideration.

In a general sense, the cost function takes the form

$$J(\mathbf{X}) = J_o(\mathbf{X}) + J_p(\mathbf{X}) \tag{7}$$

where $\mathbf{X}$ is the control variable vector. J_o measures the distance of the model solution from the observations and J_p, referred to as penalty term, includes all the physical constraints to be imposed on the model solution.

The first term of Eq. 7 is often expressed as

$$J_o(\mathbf{X}) = \frac{1}{2}(C\mathbf{Y} - \mathbf{d})^T \mathbf{W}^{-1}(C\mathbf{Y} - \mathbf{d}) \tag{8}$$

where $\mathbf{d}$ and $\mathbf{Y}$ are the vectors containing the observations and the model variables, respectively, and C is an interpolation matrix which maps the model variables to the space and time locations of the observations. The matrix $\mathbf{W}^{-1}$ is ideally the inverse of the error covariance matrix for the observations. If $\mathbf{W}^{-1}$ is approximated by a diagonal matrix, J becomes a weighted sum of squares and the technique simply corresponds to a least-squares fit method. In practice, the value of the elements of $\mathbf{W}^{-1}$ are determined by the relative magnitude of the various model variables, their dimensional scaling and the quality of the data sets. The specific form for J used in this study is given in Section 3.6.

The addition of a penalty term J_p to the cost function can provide smoother model solutions and estimated parameters, and result in a much faster convergence of the minimization process. Indeed, Sasaki (1970) showed that the penalty term suppresses the high frequencies and wave numbers in the solution. Courtier and Talagrand (1990), Zou *et al.* (1992a, 1993) showed that the penalty term can control

the amount of spurious gravity waves. Richardson and Panchang (1992), Lardner *et al.* (1993) introduced a penalty term in the cost function in order to penalize large variations in the recovered parameters and avoid negative drag coefficients. The penalty term might also introduce prior information on the parameters to be recovered and lead to a unique solution (Carrera and Neuman, 1986a,b,c). In that case, the penalty term convexifies the cost function and renders the Hessian of the cost function with respect to the control variables to be positive definite, resulting in a unique solution.

3.4 Adjoint model

While the adjoint method is a very powerful tool for obtaining the gradient of the cost function with respect to the control variables, the most difficult aspect of this technique is the development of the adjoint model code. Several approaches have been taken in the recent years.

One of the methods consists in the derivation of the continuous adjoint equations from the model equations followed by their discretization using the same scheme as for the direct model equations. The derivation of the continuous adjoint equations can be done following various approaches: the derivation of Euler-Lagrange equations (Morse and Feshbach, 1953), the control theory (Le Dimet and Talagrand, 1986), the Lagrangian multiplier approach (Thacker and Long, 1988; Smedstad and O'Brien, 1991; Lardner and Das, 1992).

Another approach consists of deriving the adjoint code directly from the discretized model code. In that case, the adjoint code can either be built from the tangent linear model code (Talagrand, 1991; Navon, 1992a) or it can be constructed based on the use of Lagrange multipliers (Lawson *et al.*, 1995a). Both techniques have two main advantages: they reduce the complexity of the construction of the adjoint model; and, they avoid the inconsistency that can arise from the deriva-

tion of the adjoint model followed by its discretization. Since the tangent linear method (TLM) and the Lagrange multiplier (LM) techniques are straightforward to implement, a full description of those methods and a simple example of their application is given in Appendices A.1 and A.2. It should be noted that, recently, adjoint compilers have been developed (Giering, 1995). However, their application is still limited to simple models.

3.4.1 Tangent linear model technique

The derivation of the adjoint model code from the tangent linear model code and the computation of the gradient of the cost function are based on the following principle. Let $\mathbf{Y} = (y_1, y_2, ..., y_m)$ represent the model variable vector, $\mathbf{X} = (x_1, x_2, ..., x_n)$ the control variable vector, and write the discretized model equations as

$$\mathbf{Y} = G(\mathbf{X}) \tag{9}$$

where G can be either a linear or a non-linear operator. If the input (control) variables are perturbed, represented by the vector $\delta\mathbf{X}$, the resulting perturbation to first order in the control variables is

$$\delta y_j = \sum_1^n \frac{\partial y_j}{\partial x_i} \delta x_i$$

or in matrix form

$$\delta\mathbf{Y} = G'(\delta\mathbf{X}),$$

where G' is the Jacobian, given by

$$G' = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots \\ & \vdots & \end{bmatrix}$$

Note that G' is the tangent linear operator of G linearized at the vicinity of $\mathbf{X}$.

Let J represent a cost function where $J = J(\mathbf{X})$. The gradient of J with respect to the control variables can be written as

$$\nabla_{\mathbf{X}} J = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots \\ & \vdots & \end{bmatrix} \begin{bmatrix} \frac{\partial J}{\partial y_1} \\ \frac{\partial J}{\partial y_2} \\ \vdots \end{bmatrix}$$

which means that

$$\nabla_{\mathbf{X}} J = (G')^* \nabla_{\mathbf{Y}} J \tag{10}$$

where $(G')^*$ is the transpose of (G') or adjoint of the tangent linear operator. Hence, the tangent linear adjoint method consists in numerically computing the gradient of the cost function with respect to the control variables using Eq. 10.

If the tangent linear operator (G') is viewed as the multiplication of a number of operators, *i.e.*,

$$G' = G'_N G'_{N-1} ... G'_2 G'_1$$

the equivalent adjoint operator is then

$$G'^* = G'^*_1 G'^*_2 ... G'^*_{N-1} G'^*_N \tag{11}$$

Each operator in Eq. 11 can be a subroutine, a DO loop, or a line in the FORTRAN code. The essence of the adjoint method is then to systematically perform computations like Eq. (10) for all steps of the basic code which will be in general a single line of code. Note that in the adjoint code, the input and output spaces of the direct code will be reversed. One should also note that only one integration of the adjoint code in reversed order from the direct model is then required to obtain the gradient of the cost function with respect to the control variables. From Eqs. 7, 8, 10 and 11, one indeed obtains

$$\nabla_X J = G'^*_1 G'^*_2 ... G'^*_{N-1} G'^*_N \left[\mathbf{W}^{-1} C (C\mathbf{Y} - \mathbf{d}) + \nabla_Y J_p \right] \tag{12}$$

and the expression in brackets can be interpreted as the forcing of the adjoint model. A simple example of the application of this method is presented in Appendix A.1.

3.4.2 Lagrange multiplier method

The construction of the adjoint code based on a scheme using Lagrange multipliers has been developed by Lawson *et al.* (1995a) and is based on the following principle. Consider the cost function (Eq. 8) as the last variable to be computed in the sequence of computations in the direct model, *i.e.*,

$$y_{m+1} = J(\mathbf{X}, y_1, ..., y_m).$$

The model equation (Eq. 9) can be rewritten as

$$y_1 = g_1(\mathbf{X}), \quad y_j = g_j(\mathbf{X}, y_1, y_2, \ldots, y_{j-1}), \quad \text{for } 1 < j \leq m+1 \tag{13}$$

and the Lagrange function, L, is then defined as

$$L(\mathbf{X}, \mathbf{Y}, \mathbf{\Lambda}) = y_{m+1} - \lambda_1(y_1 - g_1(\mathbf{X})) - \sum_{i=2}^{m+1} \lambda_i(y_i - g_i(\mathbf{X}, y_1, \ldots, y_{i-1})) \tag{14}$$

where $\mathbf{\Lambda} = (\lambda_1, \ldots, \lambda_{m+1})$ is the vector of Lagrange multipliers. Lawson *et al.* (1995a) showed that at a saddle point, which corresponds to a point in the space $\mathbf{Y}, \mathbf{X}, \mathbf{\Lambda}$ where all the derivatives of L vanish simultaneously, one obtains the model equations, the adjoint equations and the gradient of the cost function with respect to the control variables.

Requiring that the derivatives of Eq. 14 vanish with the Lagrange multipliers and the control variables, respectively, yields the adjoint equations

$$\lambda_{m+1} = 1 \quad \text{and} \quad \lambda_j = \sum_{i=j+1}^{m+1} \frac{\partial g_i}{\partial y_j} \lambda_i, \quad j = m, \ldots, 1 \tag{15}$$

and the gradient of the cost function with respect to the control variables

$$\frac{\partial J}{\partial x_k} = \frac{\partial L}{\partial x_k} = \sum_{j=1}^{m+1} \frac{\partial g_j}{\partial x_k} \lambda_j, \quad 1 \leq k \leq n. \tag{16}$$

The adjoint equations (Eq. 15) are then used to compute the Lagrange multipliers, which are used in Eq. 16 to compute the gradient of the cost function with respect to the control variables.

If the model code is thought as a succession of model equations (Eq. 13), the essence of the adjoint method based on Lagrange multipliers is then to systematically construct a Lagrange function as the function given in Eq. 14 for all steps of the basic code which will be in general a single line of code. Again, it is important to note that the computations in the adjoint equations are done in reverse order. A full description of the technique can be found in Lawson *et al.* (1995a). A simple example of the application of this method is presented in Appendix A.2.

3.4.3 Verification of adjoint code and gradient of the cost function

Any error introduced in the coding of the adjoint model can be devastating. Errors introduced in the coding of the adjoint model from the tangent linear code and the computation of the gradient of the cost function with respect to the control variables can be detected using the following two verification methods.

At any level of the coding of the adjoint model, the correctness can be checked based on the equality of scalar products, $< v, Au >=< A^T v, u >$. In other words, the sum of the square of the outputs of either a DO loop or a direct subroutine of the tangent linear code must be equal to the sum of the inputs of that DO loop (direct subroutine) multiplied by the corresponding outputs of the adjoint DO loop (adjoint subroutine), within the limits of computer accuracy.

A second verification of the correctness of the gradient of the cost function can be done as follows (Navon *et al.*, 1992a). Perturb the control variable vector by an amount $\alpha\mathbf{U}$ where α is a small scalar and $\mathbf{U}$ is a normalized vector, *e.g.*, $\mathbf{U} = \frac{\nabla J}{\|\nabla J\|}$. The Taylor expansion of the cost function is

$$J(\mathbf{X} + \alpha\mathbf{U}) \;=\; J(\mathbf{X}) \;+\; \alpha\mathbf{U}^T\nabla_{\mathbf{X}}J(\mathbf{X}) \;+\; O(\alpha^2). \qquad (17)$$

For

$$\phi(\alpha) = \frac{J(\mathbf{X} + \alpha\mathbf{U}) \; - \; J(\mathbf{X})}{\alpha \underline{U}^T \nabla_{\mathbf{X}} J(\mathbf{X})},$$

in the limit as α goes to 0, we have

$$\lim_{\alpha \to 0} \phi(\alpha) = 1 \tag{18}$$

When the Lagrange multiplier technique is used, the first verification cannot carried out. One can only verify the correctness of the gradient of the cost function as a check of the correctness of the adjoint coding. Lawson *et al.* (1995a) developed a verification technique similar to that mentioned above. Using the definition of the cost function (Eq. 8) with $C = I$, Eq. 18 becomes

$$\lim_{\alpha \to 0} \frac{(\mathbf{Y} - d)^T \mathbf{W}^{-1} \delta\mathbf{Y}}{\alpha (\nabla_{\mathbf{X}} J)^T \mathbf{U}} = 1 \tag{19}$$

where $\delta\mathbf{Y}$ is the change in the model output due to a change in the model input $\mathbf{X} = \alpha\mathbf{U}$. $\mathbf{U}$ denotes a direction along which the model input changes, *e.g.*, the gradient of the cost function. Note that the numerator of Eq. 19 is entirely determined from the model while the denominator is computed using the adjoint code.

Before any assimilation experiments, it is judicious to verify the correctness of the adjoint code. An example of the verification of the adjoint code of the circulation model in the case of the identical twin experiment described in Sections 3.6 and 4.1.2 is shown in Fig. 10. Note that for $10^{-10} < \alpha < 10^{-2}$, the ratio in Eq. 18 approaches 1.

3.5 Optimization techniques

As mentioned above, each integration of the direct model and the corresponding adjoint model provides the value of the cost function and its gradient with respect to the control variables. In most of the large-scale unconstrained local optimization algorithms, the gradient information is used in an iterative optimization process

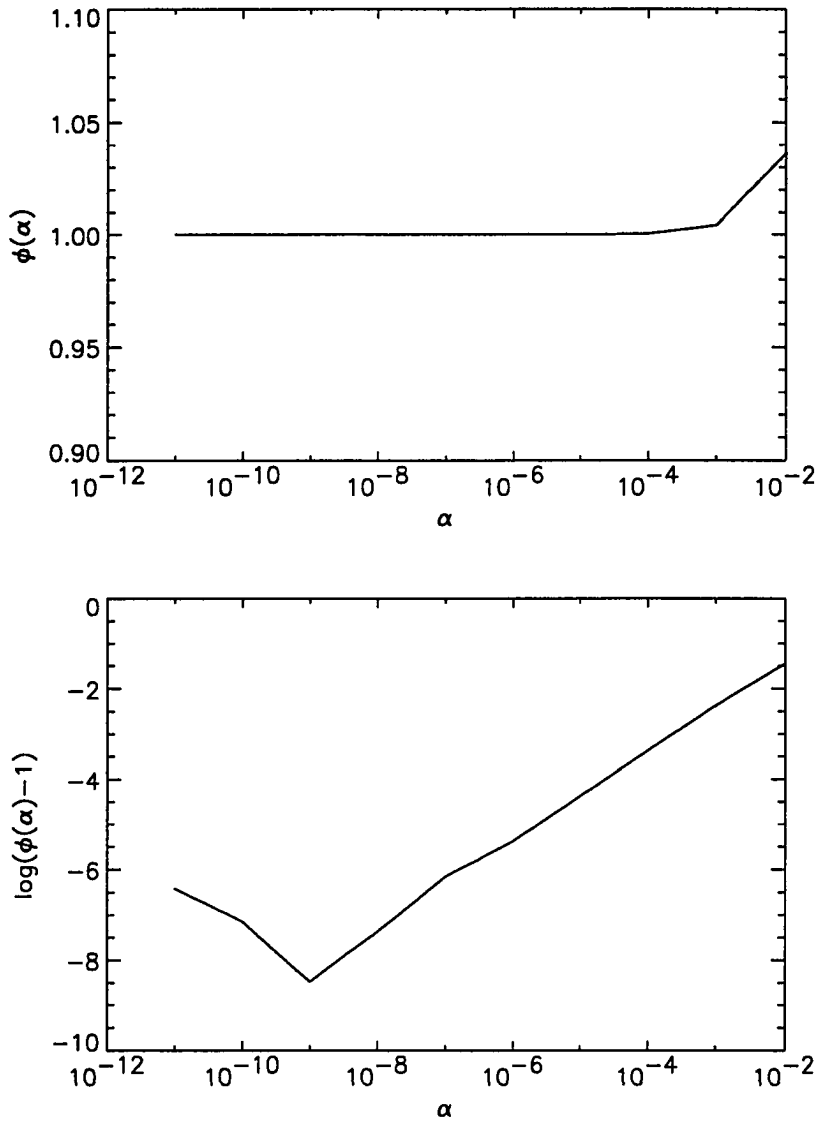

Figure 10: Verification of the gradient of the cost function using Taylor expansion.

which consists in finding the direction to adjust the control variables to minimize the misfit between model results and observations; an optimal step-size in that direction is also computed. A description of the computation of the step-size, referred to as line search, can be found in Appendix B. For linear models with a cost function as defined in Eq. 8, the number of iterations required to reach the minimum of the cost is theoretically equal to the dimension of the control variable vector **X**. However, most of the models are non-linear and many more iterations are required, which may be computationally challenging. It is therefore crucial to choose a robust and efficient minimization algorithm providing a fast convergence towards the minimum of the cost function. With that in mind, only a few optimization algorithms will be mentioned. Their characteristics are described in Appendix B.

The simplest minimization method is the steepest-descent technique where the direction is taken as minus the gradient of the function being minimized. For large problems, like in oceanography and meteorology, this technique has a very slow rate of convergence and is therefore not very efficient. Several algorithms have been developed based on the descent direction method, designed to accelerate the convergence as well as to reduce the memory required to store the matrices used to compute the direction. The most useful algorithms are: i) conjugate-gradient (C-G) methods (Fletcher, 1987; Gill *et al.*, 1981, Navon and Legler, 1987), ii) Newton and truncated Newton methods (O'Leary, 1982; Toint, 1981; Nash, 1984a,b; Schlick and Fogelson, 1992a,b; Wang *et al.*, 1992; Wang, 1993; Wang *et al.*, 1995), iii) quasi-Newton methods (Shanno, 1978; Nocedal, 1980), iv) limited memory quasi-Newton methods (Liu and Nocedal, 1989; Gilbert and Lemaréchal, 1989). The numerical implementations of these algorithms are now available from various libraries.

Based on studies by Gilbert and Lemaréchal (1989) and Zou *et al.* (1993a), the limited-memory quasi-Newton method based on Nocedal's proposal is a robust method requiring the fewest iterations and storage memory. The version N1QN3

developed by Gilbert and Lemaréchal (1989) has therefore been used in our study. This procedure computes the descent direction from an approximation of the Hessian matrix (see Appendix B.2.2) which is updated at every iteration during the minimization using the past history. The number of updates is a user defined parameter. The step-size is computed following an algorithm developed by Lemaréchal (1981) and satisfies Wolfe's conditions. A description of the method can be found in Appendix B.2.

3.6 Data assimilation in Chesapeake Bay

The main influences in the tidal and wind-driven circulation in the Chesapeake Bay are the bottom stress and the wind stress. In the past, wind stress was computed from wind speed and direction data collected at major airports along the west coast of the Chesapeake Bay. The most difficult task is then to relate wind data collected over land to the winds over the water and get an estimate of the wind stress. For example, in order to produce monthly average 1-degree pseudostress vector fields over the Indian Ocean and circumvent this difficulty, Legler *et al.* (1989) and Legler and Navon (1991) used a variational approach to develop an objective analysis technique of meteorological data. The bottom drag coefficient, needed to compute the bottom stress, is a major part of the tidal circulation, and is poorly known for the Chesapeake Bay. Hence, the main goal of this research is to use sea level measurements from tide gauges to estimate the wind forcing, bottom drag coefficient in order to get the best representation of the circulation in the Bay for the period under consideration. In addition, as mentioned in Section 2, the exchange between the open ocean and the Chesapeake Bay is dynamically important. It has been shown that the wind driven circulation with period of a few days can be remotely generated from the open ocean (Wang, 1979a). It is therefore desired to estimate open boundary conditions at the Bay mouth. The parameters to be optimally esti-

mated using variational data assimilation are then the two components of the wind stress, the bottom drag coefficient (more specifically, the inverse of the Manning's roughness [1/n] and the exponent α), and the surface elevation at the Bay mouth.

The assimilation technique to be used is the variational adjoint method using the MU-model with the adjoint code written directly from the MU-model FORTRAN code following the tangent linear technique (Section 3.4.1 and Appendix A.1). The optimization routine is the N1QN3 routine developed by Gilbert and Lemaréchal (1989) (Appendix B.2.2). The first step in the variational method consists of defining a cost function which is then minimized with respect to the control variables. The cost function (Eq. 8), which measures the difference between the model results and the observations, is chosen to be

$$J_o = \frac{1}{2}\sum_{t,s}(\eta(t,\vec{s}) - \hat{\eta}(t,\vec{s}))^2 \tag{20}$$

where t and $\vec{s}$ are the time and location of data. $\hat{\eta}$ and η are the observed sea surface elevation and model-derived values interpolated to the location of the observations. The interpolation scheme is simple bilinear interpolation. The weight matrix in Eq. 8 is taken to be the identity matrix so that elevation observations are considered to have equal importance at all the tide gauges in the Bay. A penalty term, J_p, similar to the term taken by Lardner *et al.* (1993) and Richardson and Panchang (1992), is added when wind forcing and boundary conditions are recovered. This term was found to be necessary to ensure smoothness of the recovered wind field and elevation at the Bay mouth (see Section 4.1). It is defined as

$$J_p = \frac{1}{2}\beta_1 \sum_{k}^{N-1}\left[(\tau^x_{w,k+1} - \tau^x_{w,k})^2 + (\tau^y_{w,k+1} - \tau^y_{w,k})^2\right] + \frac{1}{2}\beta_2 \sum_{t}\sum_{j}^{M-1}(\eta_{b,j+1}(t) - \eta_{b,j}(t))^2 \tag{21}$$

where τ^x_w and τ^y_w are the components of the wind stress and η_b the surface elevation at the Bay mouth. N and M are the total number of estimated values for the wind stress and the number of grid points along the open boundary, respectively, and

t is the time at which the boundary elevation is recovered. Several experiments showed that the penalty coefficients $\beta_1 = 0.1$ and $\beta_2 = 0.01$ give the best results; the oscillations in the recovered values are greatly reduced while the structure is preserved.

Since the minimization algorithm, such as the limited memory quasi-Newton method used in this study, uses an approximation of the Hessian matrix, the relative sizes of the parameters to recover have an effect on the rate of convergence (Navon *et al.*, 1992; Seiler, 1993). It is important to choose the right scaling; otherwise the optimization may yield senseless parameters or fail to converge. A general form of the scaling procedure is

$$\mathbf{X} = \mathbf{S}\mathbf{X}^s$$
$$\mathbf{g}^s = \mathbf{S}\mathbf{g}$$
$$\mathbf{H}^s = \mathbf{S}\mathbf{H}\mathbf{S}$$

where $\mathbf{S}$ is a diagonal matrix containing the scaling factors, and $\mathbf{X}$, $\mathbf{g}$, $\mathbf{H}$ are the control variable vector, the gradient vector and the Hessian matrix, respectively, and s indicates a scaled matrix or vector. Here, only the inverse of the Manning's roughness is two orders of magnitude larger than the other control variables. Dividing that parameter by 100 would then reduce it to the same order of magnitude as the other control variables and thus should accelerate the convergence. However, it was found that scaling the Manning's roughness damaged the recovery when data are sparse (Section 4.1). Therefore, no scaling was used for any of the experiments. More advanced scaling procedures can be found in Gill *et al.* (1981) but these have not been investigated.

The model domain and grid are represented in Fig. 11. The domain includes not only the main stem but also the tributaries. The grid size is 1′ in latitude ($\Delta y = 1.8$ km) and 1.25′ in longitude ($\Delta x = 2.0$ km) giving grid dimensions of 168x68. The depths at the grid points are interpolated from the NOS 15-second grid data set (see Section 2). In order to keep the problem simple, the river outflows are not taken

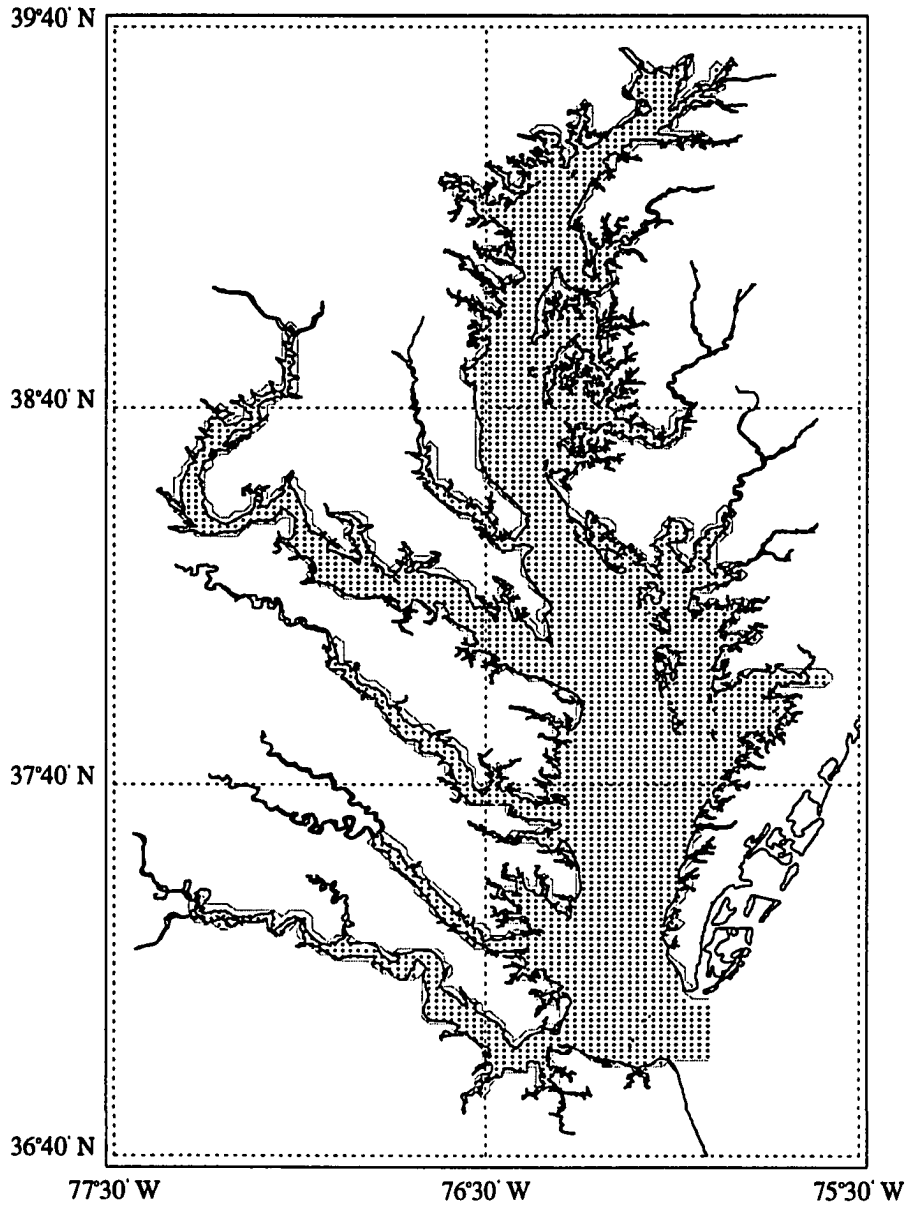

Figure 11: Model domain (dotted line) and grid. The dots represent grid points.

into account in the simulations, and closed boundary conditions are chosen at the head of the rivers. Therefore, the simulations correspond to periods when the river discharges are minimal and can be neglected. At the mouth of the Bay, the open boundary conditions are either imposed by the tidal forcing defined in Eq. 6 (tidal circulation experiment) or assimilated (identical twin experiments and wind-driven circulation experiments). Based on several runs of the direct model, a time step of $\Delta t = 10$ min seems to be appropriate to represent the circulation in the Bay. Note that the adjoint model was run with the same grid spacing, bathymetry and time step as the forward model.

The recovery experiments presented in this study are divided into three categories: identical twin, tidal circulation and wind-driven circulation experiments. They essentially differ by the nature and density of available observations for assimilation as well as the kind of control parameters that are recovered. In identical twin experiments, the observations are hourly sea surface elevations generated by the circulation model with tidal forcing and either a northeasterly or a southwesterly wind. Those observations are assimilated either at every grid point of the model domain or at ten locations corresponding to the permanent tide gauge stations in the Bay (Fig. 12). In tidal circulation experiments, the observations are hourly tidal elevations at the ten permanent tide gauge stations, reconstructed from the harmonic constants of five major constituents of the tidal signal in the Bay (M_2, S_2, N_2, K_1, O_1) using Eq. 6. In the wind-driven circulation experiment, the observations are the hourly sea surface elevation measured at the ten permanent tide gauge stations. A detailed description of the observations used in the different experiments is given in Section 4.

In all the experiments, hourly surface elevations are assimilated for a period of 24 hours. The choice of the assimilation window was dictated by the time scale of the tidal and wind-driven circulation in the Bay, the limit of the computer capacity,

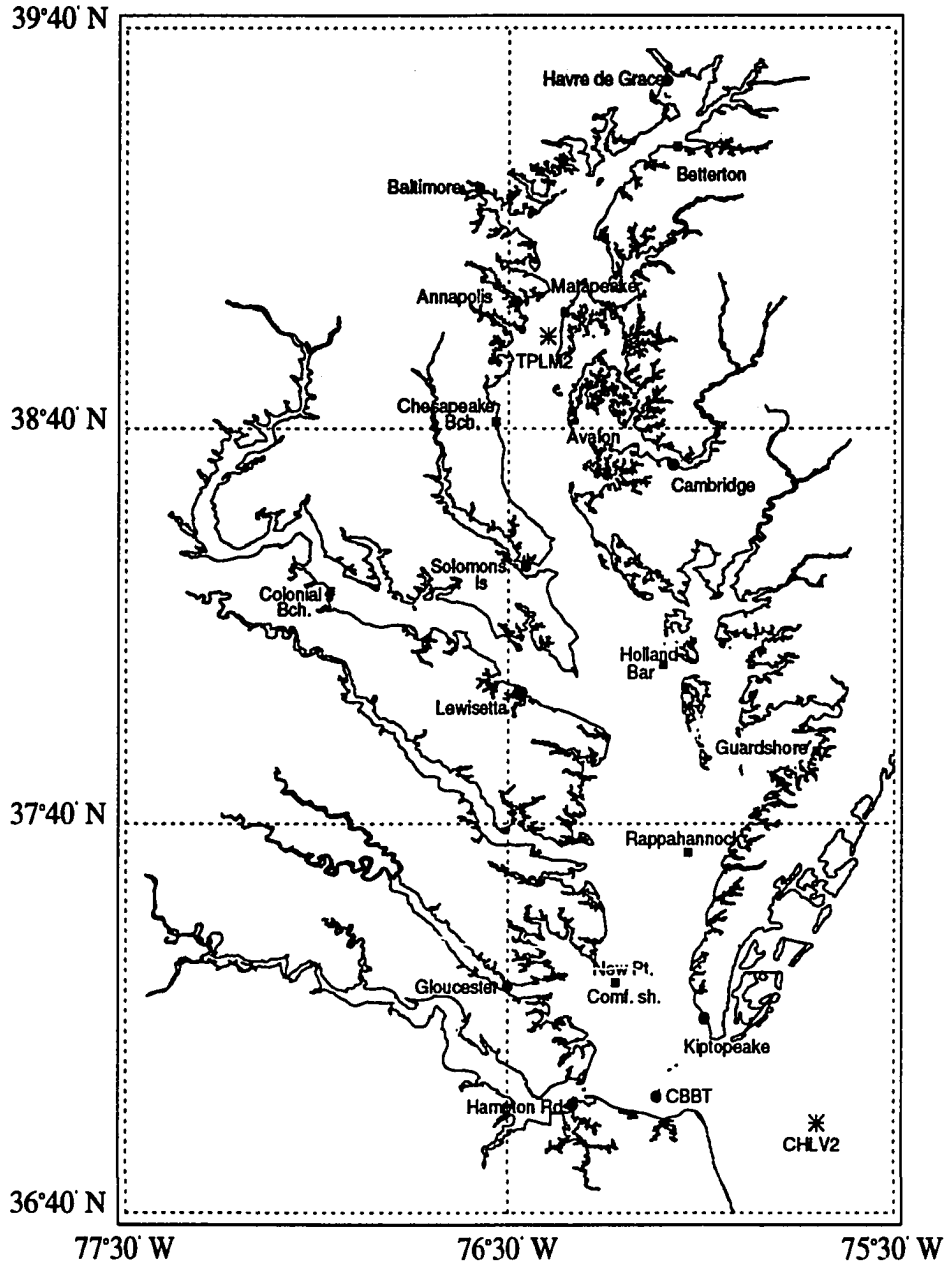

Figure 12: Locations of the tide gauge stations and buoys. The circles indicate permanent gauges and the squares indicate the stations used for comparison between modeled and observed elevations. The stars represent buoys. TPLM2 and CHLV2 designate the buoy at Thomas Point and Chesapeake Light Tower, respectively.

and the validity of the tangent linear model approximation (Li *et al.*, 1993). Fisher (1986) showed that one complete wavelength of the semi-diurnal tide is contained within the Bay. Wang (1979a) found that seiche oscillations at period of a few days were driven by the local longitudinal winds. An assimilation window of 24 hours is therefore appropriate to estimate the main control processes, *i.e.*, bottom stress and wind stress, for the tidal and wind-driven circulation studies. Computer capacity was the second limiting factor for the length of the assimilation window. Every model variable at every time step has to be stored in order to run the adjoint model (see Appendix A.1). Furthermore, using the ADI scheme defined in Section 3.2, the model variables have to be stored at the half time-step, *i.e.*, every five minutes. An assimilation run with a 24 hour window and recovery of all the control variables of the circulation model typically required over 180 Megabytes of memory and approximatively 10 hours of CPU time on an IBM RS-6000/590. An increase of the assimilation window does not significantly improve the resolution of the physical processes under study but would have required a larger memory capacity and would have increased the computer time for each run. It was verified that the tangent linear model correctly approximated the non-linear model in the 24 hour assimilation window.

In the identical twin experiments and wind-driven circulation simulation, the bottom drag coefficient parameters (inverse of the Manning's roughness (1/n) and the exponent α (Eq. 5)), the boundary elevations at the Bay mouth and the wind stress components are estimated during the assimilation process. Since the wind over the Bay changes over a period of two to three days, the wind stress components are taken constant during the 24 hour assimilation. The wind stress components are estimated at four locations in the north-south direction while they are kept uniform in the east-west direction. They are then linearly interpolated in space to get an estimate at every grid point. The first location where wind stress components are

evaluated corresponds to the Chesapeake Bay Bridge Tunnel (CBBT) station and the other three locations are taken 80 km apart in the north-south direction. The sea surface elevation is recovered at every grid point along the open boundary and at every hour. Since these elevation values are required at every half time-step (every 5 min), the hourly values are linearly interpolated in time.

For the tidal experiments, only the bottom drag coefficient parameters are estimated. Since the tidal signal is well known at the Bay mouth, there was no need to recover the boundary elevations during the assimilation process. Instead, the open boundary condition at the Bay mouth is prescribed by the tidal forcing based on the five major constituents and defined by Eq. 6 when the phases and amplitudes of the constituents are taken at the Chesapeake Bay Bridge Tunnel. The total number of parameters to be estimated, then, varies from 2 for the tidal circulation experiment to 235 for the wind-driven circulation and identical twin experiments.

For all experiments, the main question is: does the new value of the parameters really improve the model results? To evaluate quantitatively the performance of the data assimilation, three quantities were computed: the root mean square (rms) error defined as

$$rms = \left\{ \frac{1}{N} \sum_{i=1}^{N} (\eta_i - \hat{\eta}_i)^2 \right\}^{1/2}, \tag{22}$$

the relative average error (E) defined as

$$E = 100\% \frac{\sum_{i=1}^{N} (\eta_i - \hat{\eta}_i)^2}{\sum_{i=1}^{N} (|\eta_i - \bar{\hat{\eta}}|^2 + |\hat{\eta}_i - \bar{\hat{\eta}}|^2)} \tag{23}$$

and the correlation coefficient given by

$$r = \frac{\sum_{i=1}^{N} (\eta_i - \bar{\eta})(\hat{\eta}_i - \bar{\hat{\eta}})}{\left(\sum_{i=1}^{N} (\eta_i - \bar{\eta})^2 \sum_{i=1}^{N} (\hat{\eta}_i - \bar{\hat{\eta}})^2 \right)^{1/2}} \tag{24}$$

where η_i and $\hat{\eta}_i$ represent the time series of the modeled and observed elevation at some location, respectively, and "over bar" denotes time mean values. N is equal to 24, the number of observations for the 24 hour assimilation window. The root

mean square and relative average error give a measure of the difference in amplitude between modeled and observed elevations while the correlation coefficient gives a measure of the phase shift. It is important to consider all three of these quantities when evaluating the success of the data assimilation. Indeed, a small relative error with a small correlation indicates a phase shift between the observations and the recovery, an indication of poor performance of the data assimilation procedure.

4 Results

Optimal parameter estimation in the Chesapeake Bay is investigated with three types of experiments: identical twin experiments, experiments focusing on tidal circulation, and experiments focusing on wind-driven circulation.

Since it is important to check the performance of the data assimilation procedure, identical twin experiments with data generated by the circulation model itself are considered first. This corresponds to the best possible situation since the data are not contaminated with observational errors and contain the same dynamics as the circulation model used in the assimilation procedure.

As shown in Section 2, two types of circulations, *i.e.*, tidal and wind-driven circulations, are dominant in the Chesapeake Bay. The emphasis of the second set of experiments is on tidal circulation. The assimilated observations are hourly predicted sea surface elevations reconstructed at ten permanent tide gauge locations from the harmonic constants determined by Fisher (1986). The focus of these experiments is on the estimate of the bottom drag coefficient needed to compute the bottom stress.

The third set of experiments focuses on the wind-driven circulation and emphasis is on the estimate of the bottom and surface stresses. The observations used in this case are hourly sea surface elevations measured at ten permanent tide gauges. For all three types of experiments, the observations are taken in November when the stratification in the entire Bay is weak. Depending on the availability of observations, either the year 1983 or 1990 has been considered. A detailed description of

the observations and the results of the assimilation for the mentioned experiments are presented in the following sections.

4.1 Identical twin experiments

Identical twin experiments use observations generated by the circulation model and focus on the feasibility of data assimilation in the Chesapeake Bay and on the performance of the technique. In order to address those questions, the set of identical twin experiments is divided in four classes. In the first experiment, the observations are the hourly model generated surface elevation with tidal forcing and a north-easterly wind. The wind speed varies between 6.5 and 7.5 m/s over the entire Bay. Data are assimilated at every grid point of the model domain. The second experiment differs from the first one by the density of available observations. Only hourly model generated elevations at locations corresponding to the permanent tide gauge stations are assimilated. The third experiment repeats the second experiment but the model is forced with a southwesterly wind. Finally, the fourth experiment shows the importance of the penalty term in the cost function (Section 3.6) and the effect of the scaling of the inverse of the Manning's roughness (Section 3.6) on the recovery. In all the experiments, the bottom drag coefficient, the hourly boundary elevation and the four wind stress components (Section 3.6) are estimated during the assimilation process. The assimilated observations and results of the four experiments are presented in the following two subsections.

4.1.1 Model derived observations

Two time series of observations, *i.e.*, sea surface elevations, are generated using the circulation model described in Section 3.2 with predefined values for the control variables (Table 2 and Section 3.6), bottom drag coefficient parameters, four values of the wind stress components and hourly elevation at the Bay mouth. The wind

Parameter	Units	Value
α		0.15
1/n	$m^{-\alpha}$	65.00
$\tau^w_{x,1}$	Nm^{-2}	$\pm$ 0.03
$\tau^w_{x,2}$	Nm^{-2}	$\pm$ 0.06
$\tau^w_{x,3}$	Nm^{-2}	$\pm$ 0.04
$\tau^w_{x,4}$	Nm^{-2}	$\pm$ 0.07
$\tau^w_{y,1}$	Nm^{-2}	$\pm$ 0.10
$\tau^w_{y,2}$	Nm^{-2}	$\pm$ 0.08
$\tau^w_{y,3}$	Nm^{-2}	$\pm$ 0.09
$\tau^w_{y,4}$	Nm^{-2}	$\pm$ 0.08

Table 2: Summary of the parameter values used to generate the set of observations for the identical twin experiments.

stress field, representing two typical winds in the Chesapeake Bay, a northeasterly and a southwesterly wind, was generated as described in Section 3.6. Note that in the northeasterly (southwesterly) wind case, both components of the wind stress are negative (positive) (Table 2). In order to create observations corresponding to the same period of time as for real data assimilation, the hourly boundary elevations were taken equal to the predicted elevation at the Chesapeake Bay Bridge Tunnel (CBBT) for November 1, 1983. A slope of 25 cm from the southern to the northern end of the Bay mouth was added to simulate the spatial variation of the surface elevation across the Bay mouth due to the wind and the spatial variation of the tide. The hourly surface elevations were then linearly interpolated in time in order to get a value at every half time step when boundary conditions are required.

The circulation model was first spun up for two days with only tidal forcing in order to minimize the effect of the initial conditions. The model was then run for 24 hours with wind. This simulation corresponds to the case of weak wind followed by a frontal passage when wind speeds are larger. The modeled output were then subsampled to produce the hourly surface elevations which are used as observations in the identical twin experiments.

4.1.2 Recovery

As mentioned in Section 3, variational method is an iterative process leading to the best estimate of the control variables which minimize a cost function. The recovery process then starts with a first guess for all the control variables. For all the identical twin experiments, no wind, hourly predicted elevations at the Bay mouth, and a constant drag coefficient computed with $\alpha = 0$ and $1/n = 50$ were chosen as a first guess. As an indication of the discrepancy between the modeled elevation with the first guess control variables and the observations, the relative average error over the 24 hours of assimilation was computed at every grid point

using Eq. 23. Maps of the relative average error are given in Figs. 13 and 14 for the northeasterly and the southwesterly wind case, respectively. As expected with a northeasterly wind (Fig. 13), the highest error is on the western side of the Bay. It varies from a few percents to 40% in the main stem while a maximum error of 65% is found in the tributaries (white areas in Fig. 13). It is also clearly seen that the narrow portion of the upper Bay is less affected by a northeasterly wind than the rest of the Bay. In the case of a southwesterly wind (Fig. 14), the highest error is found in the upper Bay and is much higher than with a northeasterly wind. Note that the relative average error pattern found for those experiments would not necessarily be found with real observations. For instance, the boundary condition at the Bay mouth mainly acts as a tidal forcing since the slope added to the predicted elevation is constant in time. Therefore, inflow and outflow due to the wind are not well represented. This might create higher or lower elevations in some part of the Bay than in a real situation.

• *Experiment 1: Northeasterly wind and observations everywhere*

The first assimilation experiment corresponds to the best situation when observations are available at every grid point of the model domain, *i.e.*, everywhere in the main stem and the tributaries. The recovery of the bottom drag coefficient parameters (α, 1/n) and of the four values of the wind stress components ($\tau^w_{x,i}$, $\tau^w_{y,i}$ with i=1,..,4) is shown in Fig. 15. The adjustment of the boundary elevation at the southern end of the Bay mouth after 24 hours from the beginning of the recovery day, representative of the recovery of the boundary condition, is also shown in Fig. 15. The wind stress components are adjusted within the first thirty iterations. The drag coefficient parameters remain constant until the wind stress components are completely recovered. Then, they converge towards their true values, followed by

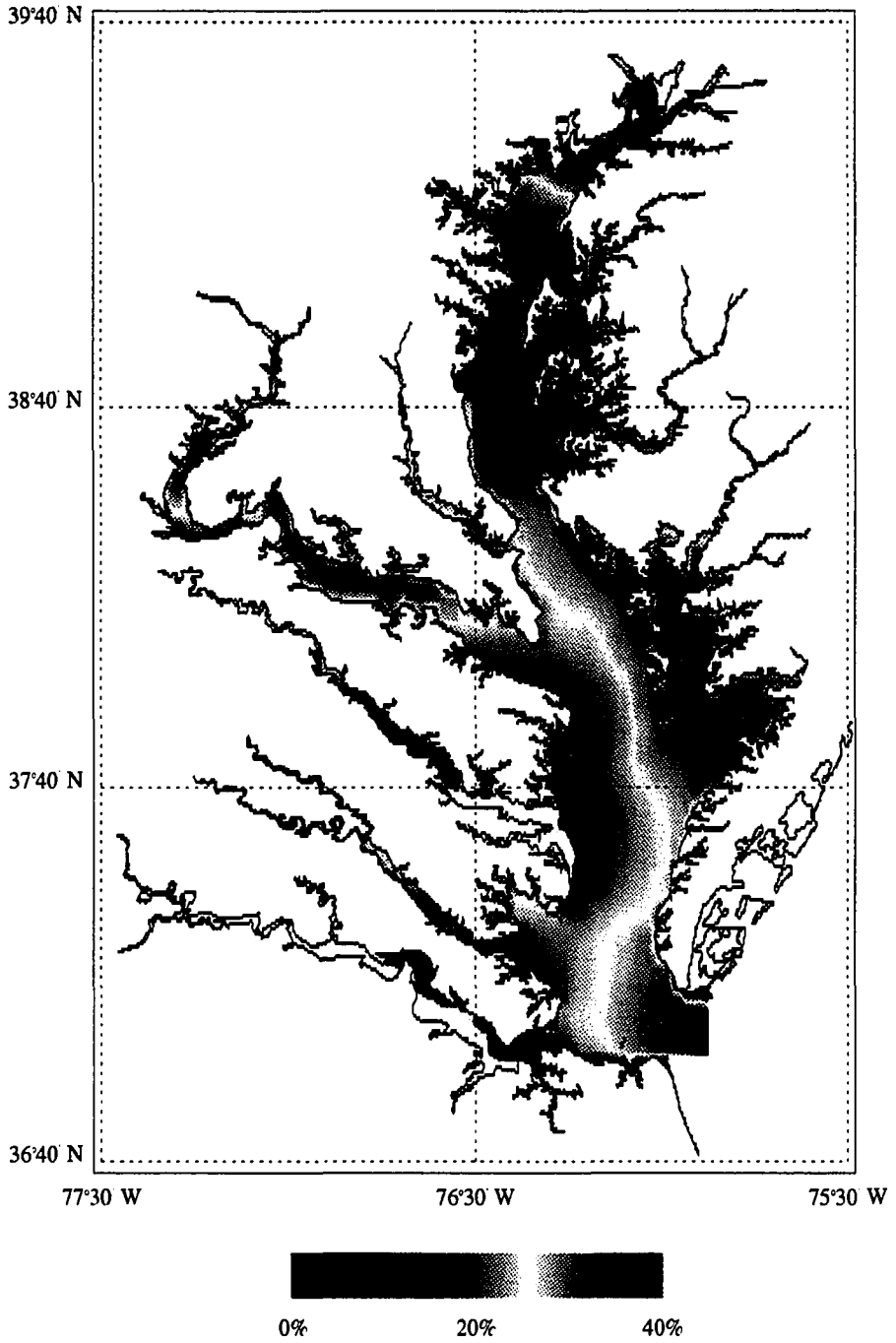

Figure 13: Map of the relative average error (%) over 24 hours between modeled elevation using the first guess control parameters and the observations for a north-easterly wind.

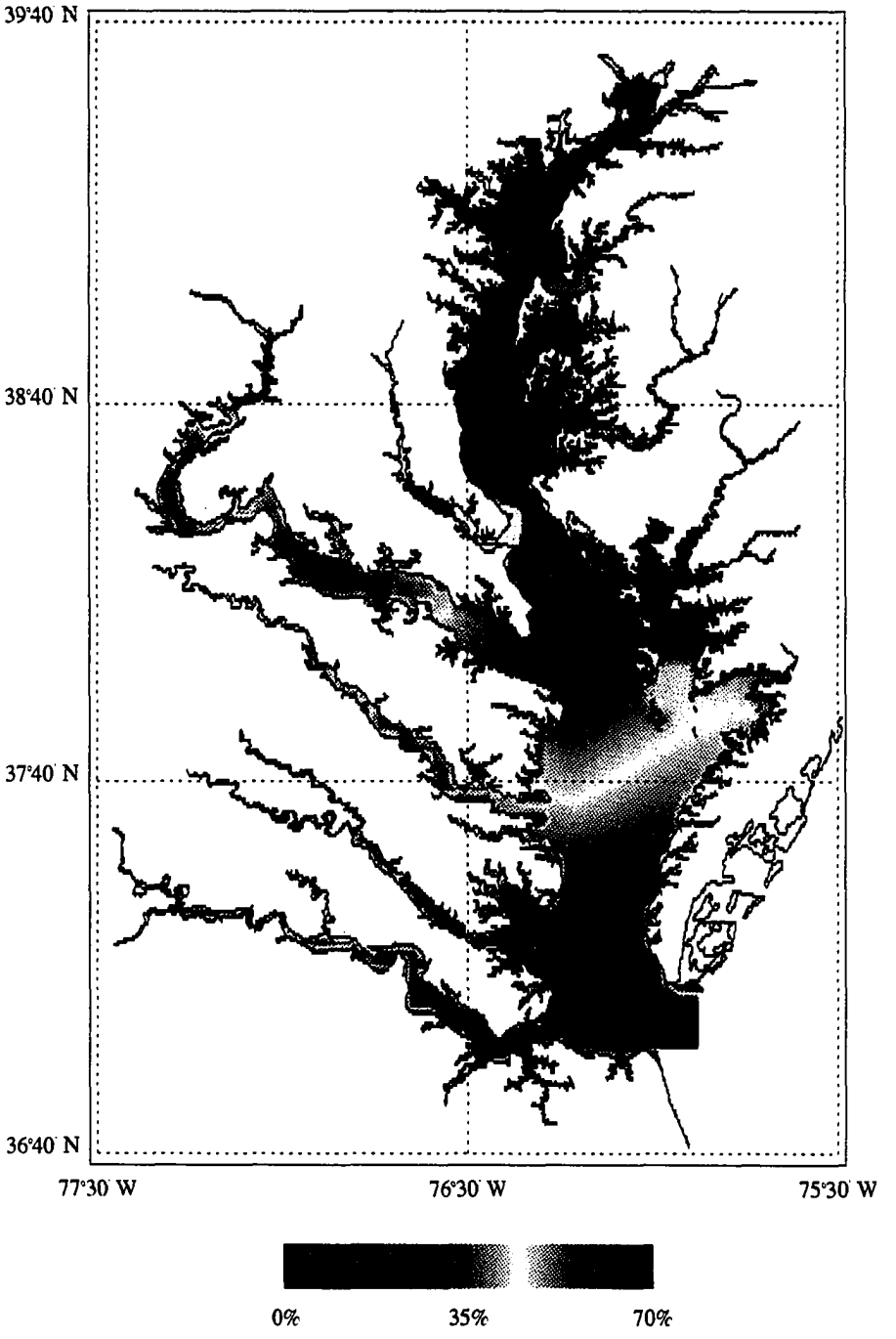

Figure 14: Map of the relative average error (%) over 24 hours between modeled elevation using the first guess control parameters and the observations for a south-westerly wind.

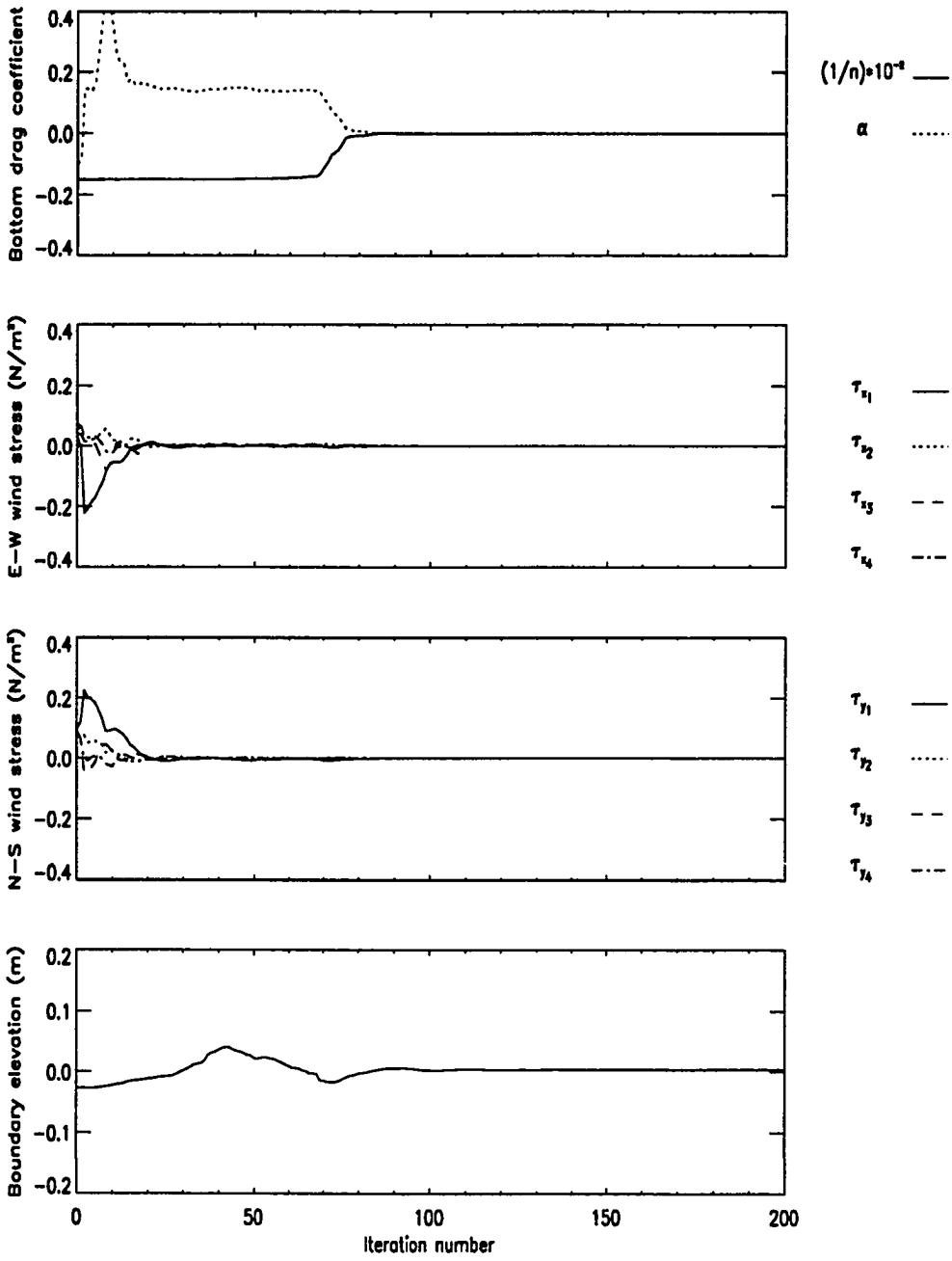

Figure 15: Difference between recovered and true values for the model parameters and boundary elevation at the southern end of the Bay mouth 24 hours after the beginning of the recovery day versus the number of minimization iterations. The data are available everywhere and a northeasterly wind is considered.

the recovery of the boundary conditions. In this case, the exact original values (four significant figures) are obtained for the bottom drag coefficient parameters and the wind stress components while a difference of less than 1 mm between estimated and true values is found for the boundary elevations. The cost function (Fig. 16) rapidly decreases during the adjustment of the wind stress components then again during adjustment of the drag coefficient parameters. The norm of the gradient of the cost function essentially follows the same pattern, except that it continues to decrease considerably after recovery of the drag coefficient and wind stress until the boundary elevations reach their true values. The pattern of adjustment of the control variables suggests that the wind stress contributes more to the data misfit for the entire Bay than the drag coefficient and the boundary condition. Indeed, after 15 iterations, when only the wind stress components are close to their true values, the relative average error between estimated and observed surface elevation (Fig. 17) shows a maximum of 1.5 % in the main stem compared to a maximum of 40 % for the first guess error (Fig. 13). Furthermore, the spatial effect of the bottom friction and the boundary condition on the surface elevation can be seen in the spatial variation of the relative average error after 15 iterations (Fig. 17). While the error is small in the main stem, the highest value is found in a small region near the Bay mouth, where the circulation is controlled by the interaction between the open ocean and the Bay and therefore is more sensitive to the boundary conditions.

• *Experiment 2: Northeasterly wind and observations at ten locations*

In real cases, observations are not available everywhere in the domain. For instance, in the Chesapeake Bay, only ten tide gauges are permanent (Fig. 12). Therefore, the previous experiment was repeated with observations available at only ten locations corresponding to the permanent tide gauge stations.

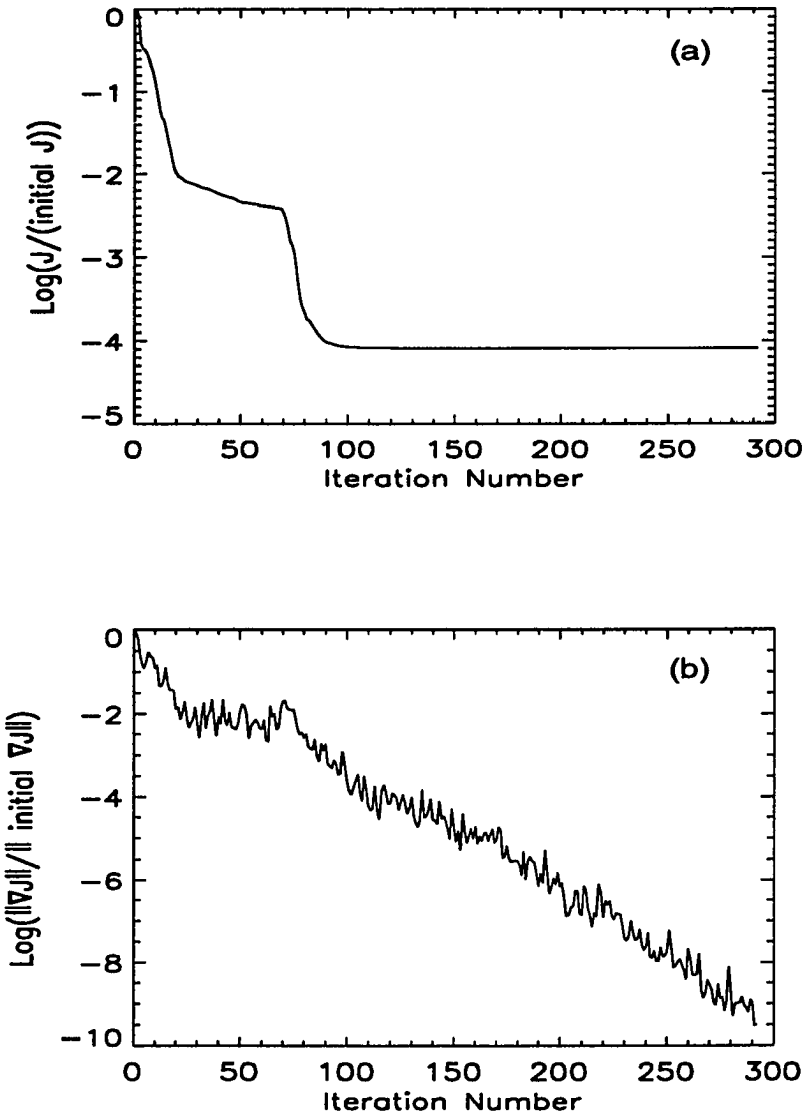

Figure 16: Logarithm of the cost function normalized by its initial value (a) and logarithm of the normalized norm of the gradient of the cost function (b) versus the number of iterations. The data are available everywhere and a northeasterly wind is considered.

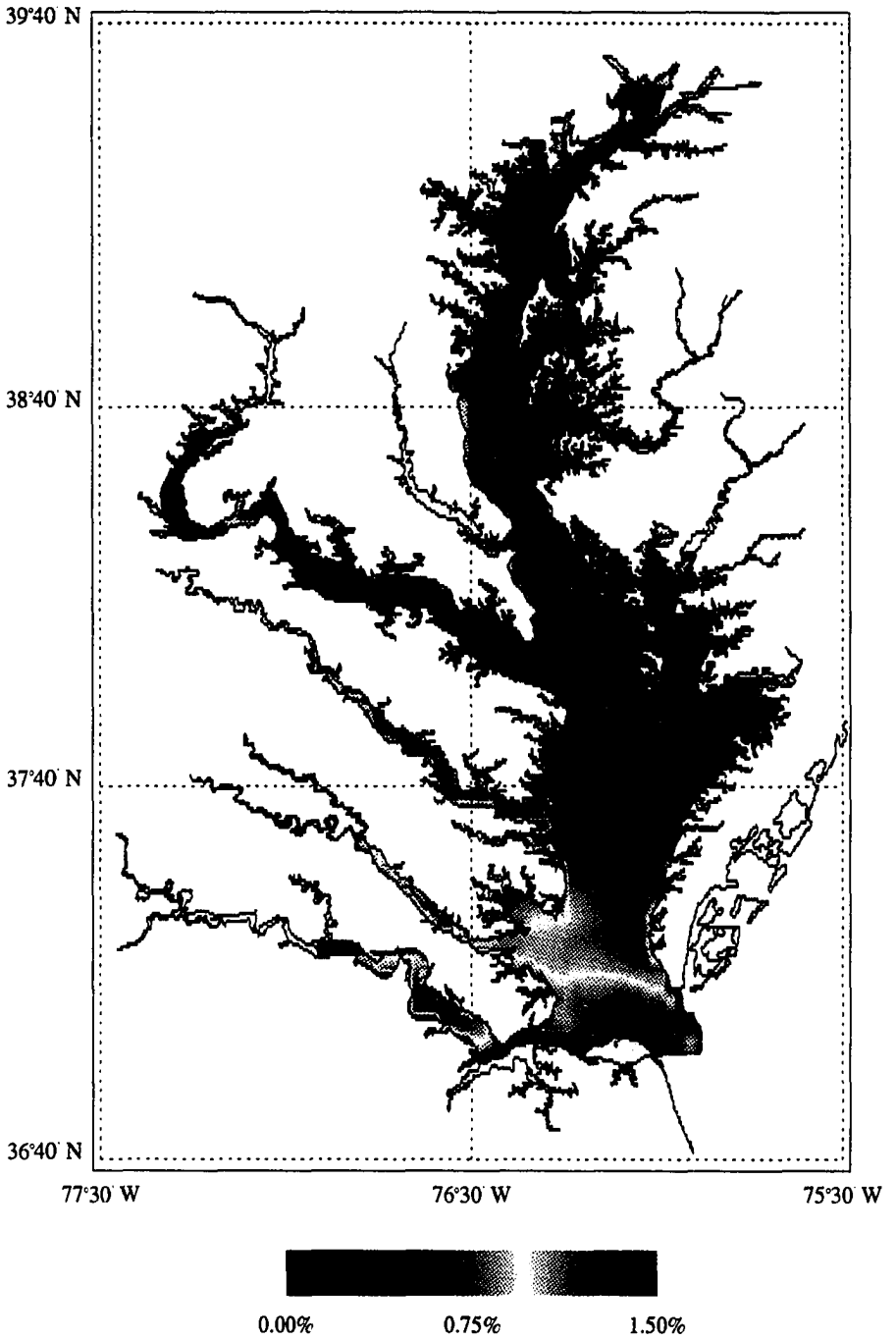

Figure 17: Map of the relative average error (%) over 24 hours between recovered and observed surface elevation after 15 iterations of the assimilation process. The data are available everywhere and a northeasterly wind is considered.

The recovery of the bottom drag coefficient parameters (α, 1/n), the four values of the wind stress components ($\tau^w_{x,i}$, $\tau^w_{y,i}$ with i=1,..,4), and the recovery of the boundary elevation at the southern end of the Bay mouth after 24 hours from the beginning of the recovery day are shown in Fig. 18. The pattern of adjustment found in the experiment with observations everywhere is also observed when data are only available at ten locations. The wind stress components is recovered within the first thirty iterations, followed by the recovery of the drag coefficient and finally the boundary elevation adjustment. While the drag coefficient and the components of the wind stress are obtained to two significant figures, the boundary conditions do not converge to their true values. The estimated boundary elevation reaches about 85% of its real value. As previously, the cost function (Fig. 19) rapidly decreases during the adjustment of the wind stress components and the drag coefficient. The norm of the gradient reaches its minimum value after approximatively 250 iterations. A map of the relative average error over 24 hours between recovered and observed elevation after 15 iterations of the assimilation process (Fig. 20) shows different patterns from when data were available everywhere (Fig. 17). Note that the color scales are different between Fig. 20 and Fig. 17. A higher error indicates that the recovery is slower. The adjustment in the middle of the Bay is mainly uniform and is faster than in the other regions of the Bay. The highest error is found at the head of the Bay (white area) where the data at the station equivalent to Havre de Grace influence the adjustment only locally. As in the previous experiment when data were available everywhere, the effect of the boundary condition on the recovery of the surface elevation is limited to a small region close to the Bay mouth. The effect of the data at the two stations corresponding to Gloucester and Hampton Road are noticeable in the process of recovery in the lower Bay. The error is indeed smaller in the surrounding region of those two stations than in other area of the lower Bay.

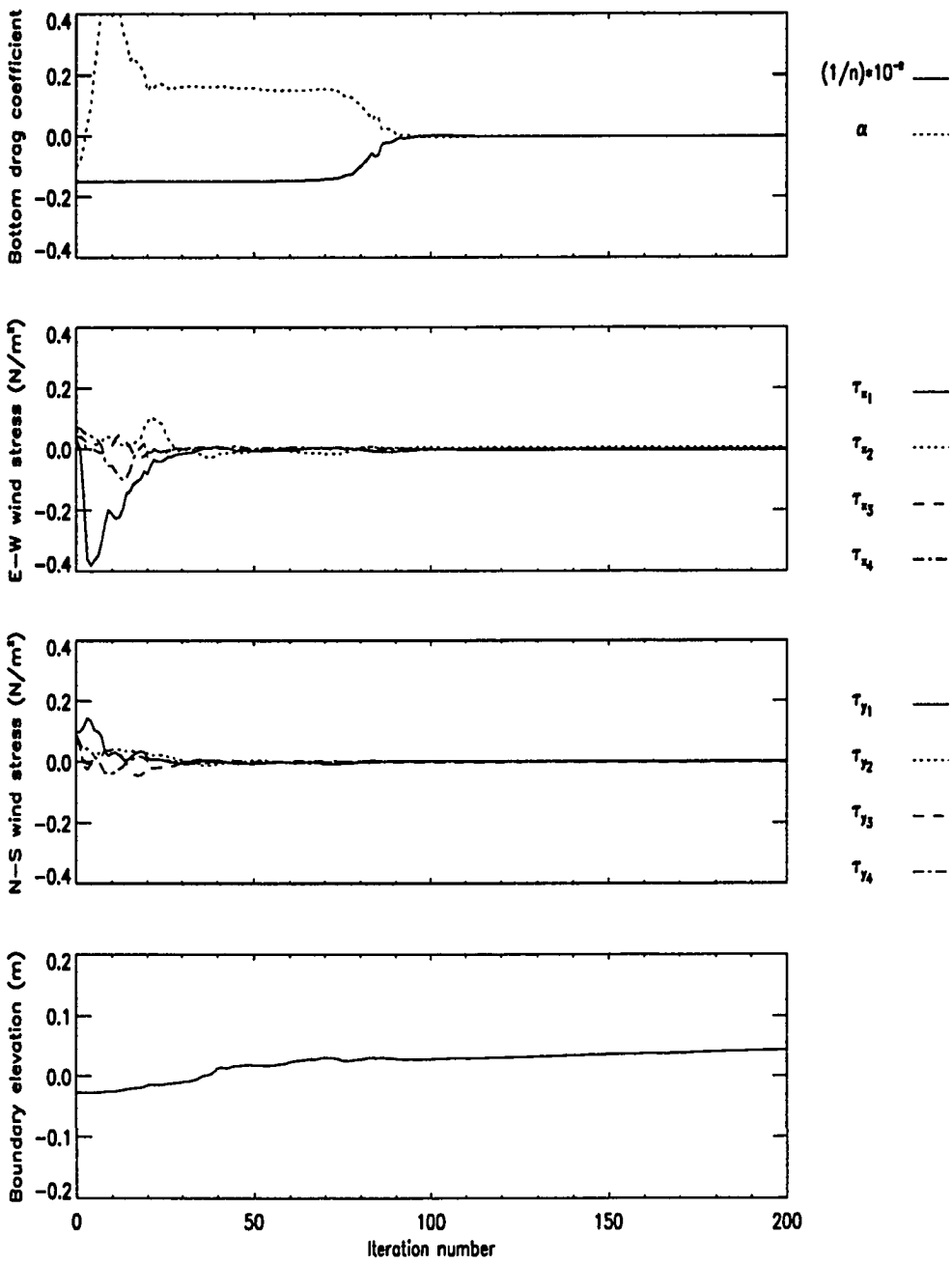

Figure 18: Difference between recovered and true values for the model parameters and boundary elevation at the southern end of the Bay mouth 24 hours after the beginning of the recovery day versus the number of minimization iterations. The data are available at ten stations and a northeasterly wind is considered.

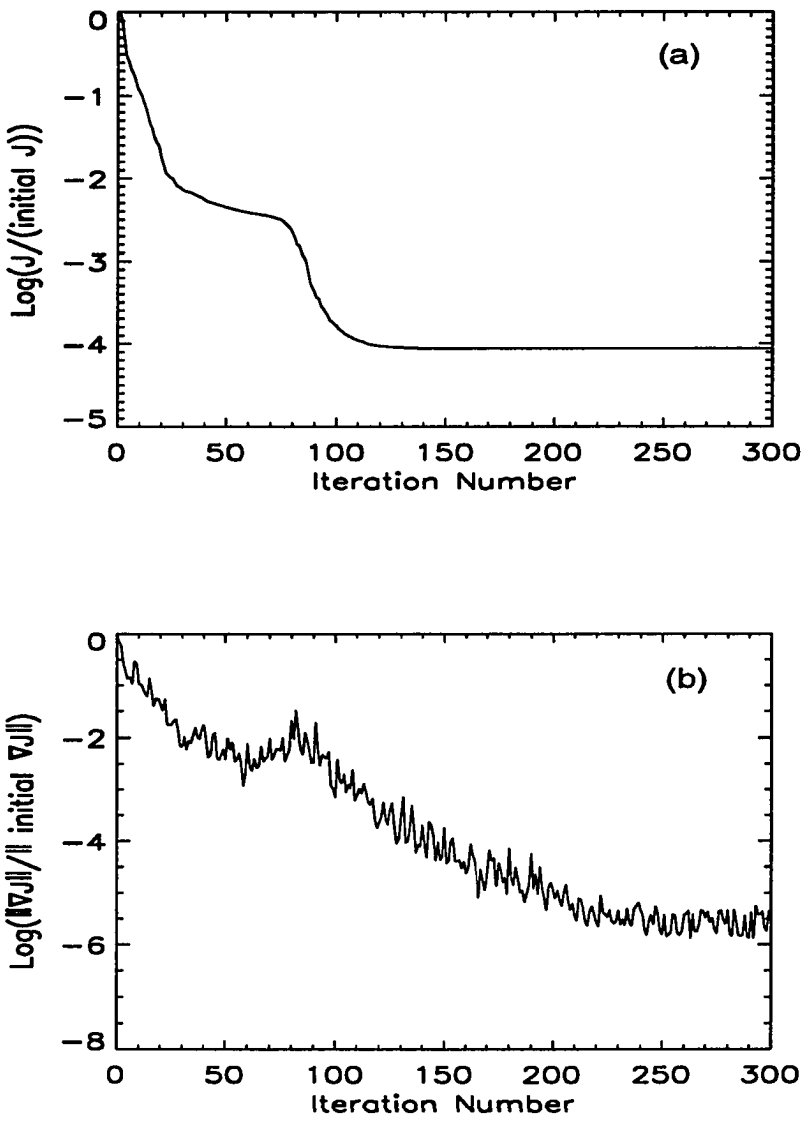

Figure 19: Logarithm of the normalized cost function (a) and normalized norm of the gradient of the cost function (b) versus the number of iterations. The data are available at ten stations and a northeasterly wind is considered.

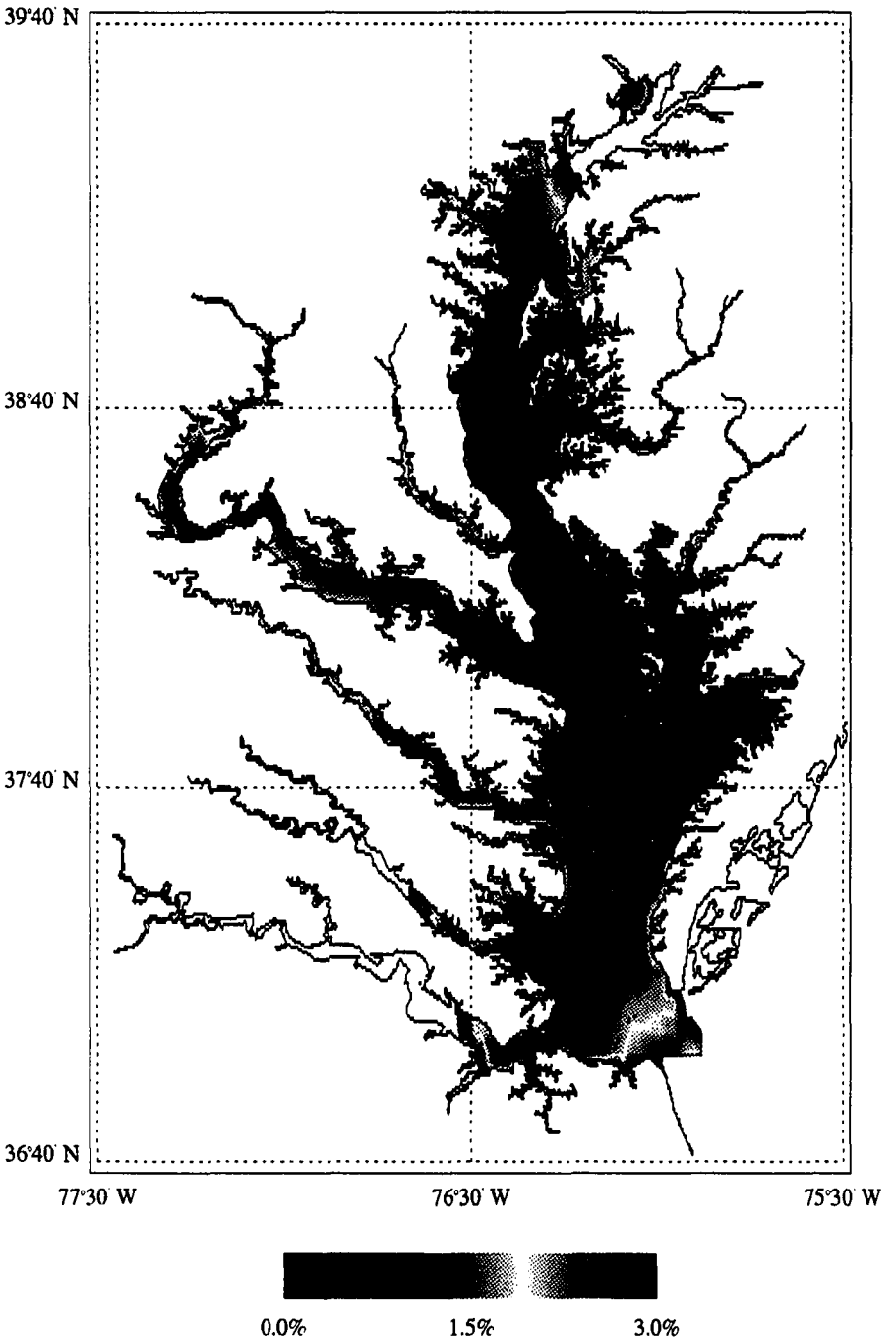

Figure 20: Map of the relative average error (%) over 24 hours between recovered and observed surface elevation after 15 iterations of the assimilation process. The data are available at ten stations and a northeasterly wind is considered.

• *Experiment 3: Southwesterly wind and observations at ten locations*

This experiment was designed to evaluate the difference in recovery when the wind blows from the southwest instead of from the northeast. A southwesterly wind is indeed typical for the Chesapeake Bay. Only observations from the ten locations corresponding to the permanent tide gauge stations were assimilated. The simulation was run the same way as the previous experiment.

While the patterns of recovery (Fig. 21) are similar to the experiment with a northeasterly wind (Fig. 18), the adjustment of the bottom drag coefficient parameters is slower for a southwesterly wind. The boundary elevation after 24 hours now reaches 87% of its real value. The better recovery of the boundary elevation is also reflected in the smaller value of the cost function and the norm of its gradient after 250 iterations (Fig. 22).

The map of the relative average error over 24 hours between estimated and observed elevations after 15 iterations of the assimilation process (Fig. 23) shows a different pattern of adjustment from the case of a northeasterly wind (Fig. 20). The elevation in the main stem seems to adjust at the same rate. The highest error is found in the middle of the Bay near the shores and at the Bay mouth which was not the case with a northeasterly wind. The effect of the boundary condition is the same for both winds. Finally, it should be pointed out that the error after 15 iterations is of the same order of magnitude in the case of a northeasterly and a southwesterly wind even though its initial value was much higher in the case of a southwesterly wind.

• *Experiment 4: Penalty term and scaling effects*

As mentioned in Section 3.6, several studies related to parameter estimation showed the necessity of adding a penalty term in the cost function. In our study, a penalty

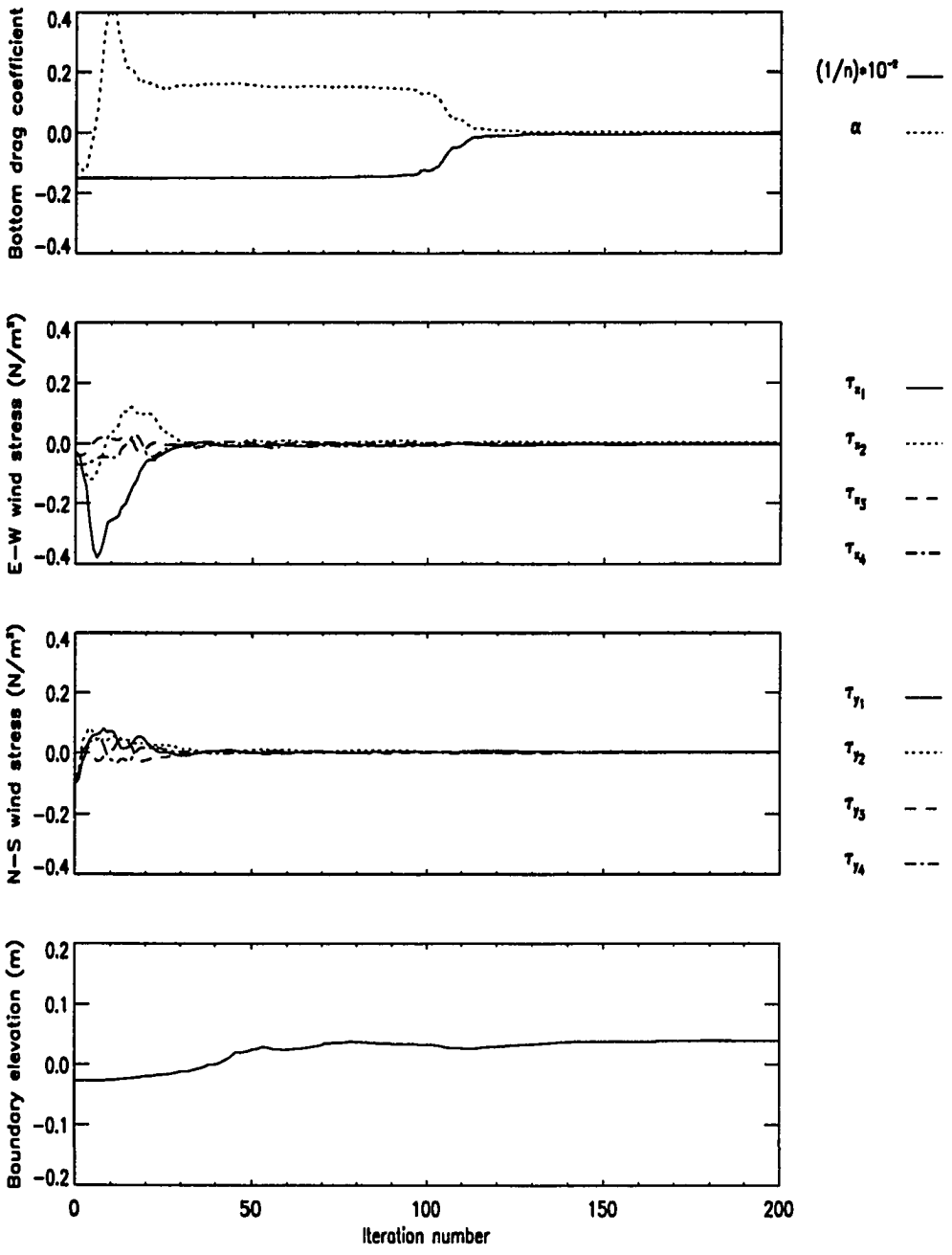

Figure 21: Difference between recovered and true values for the model parameters and boundary elevation at the southern end of the Bay mouth 24 hours after the beginning of the recovery day versus the number of minimization iterations. The data are available at ten stations and a southwesterly wind is considered.

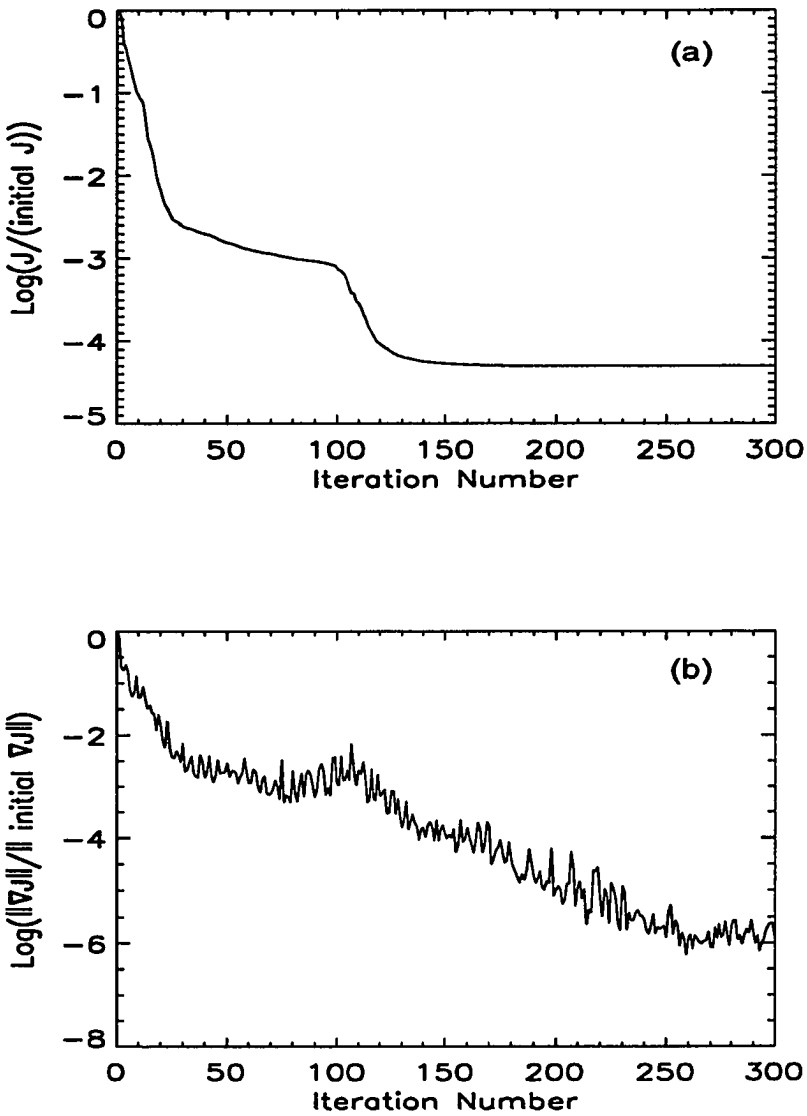

Figure 22: Logarithm of the normalized cost function (a) and normalized norm of the gradient of the cost function (b) versus the number of iterations when a southwesterly wind is blowing and data are available at ten stations.

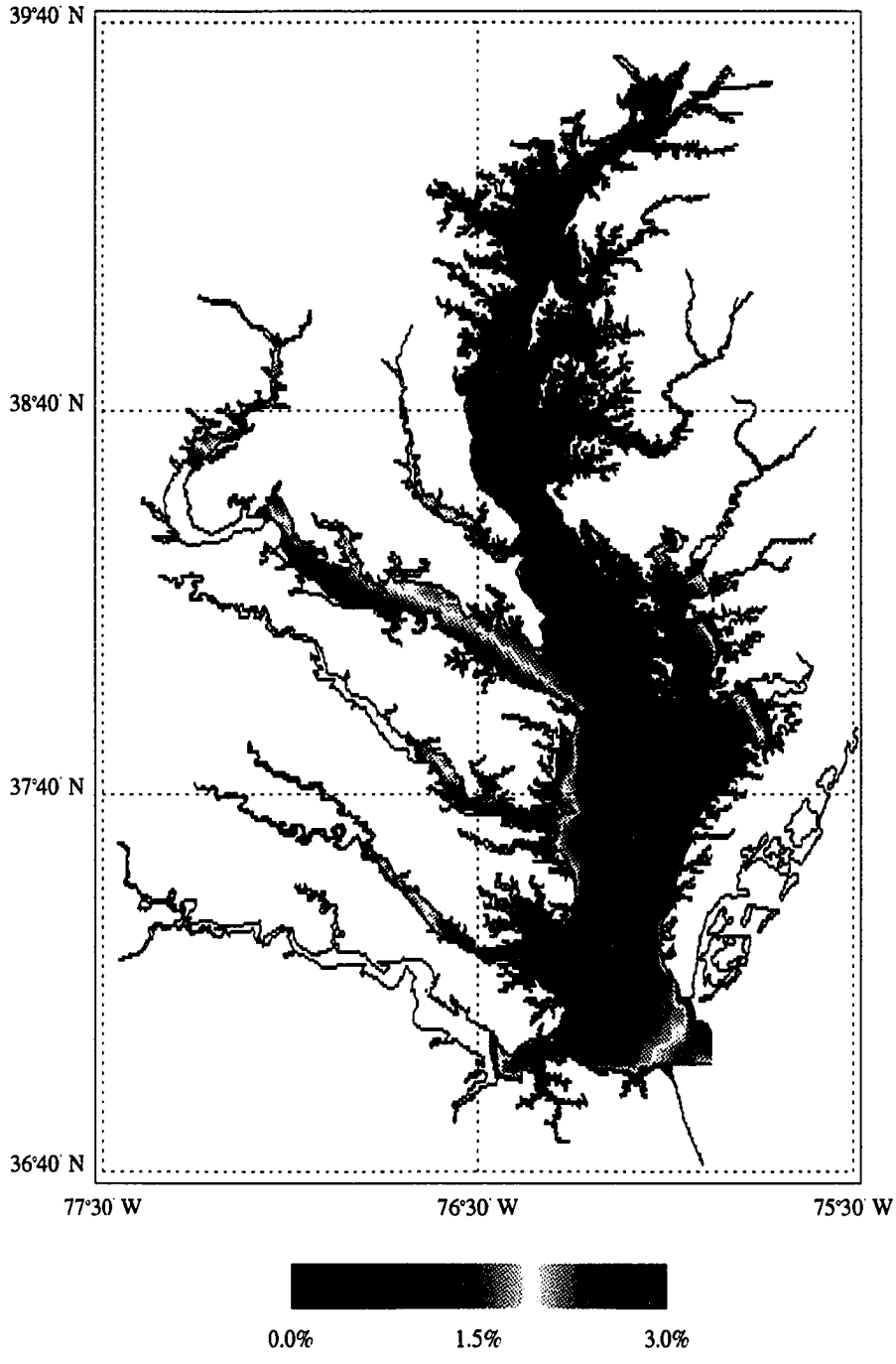

Figure 23: Map of the relative average error (%) over 24 hours between recovered and observed surface elevation after 15 iterations of the assimilation process. The data are available at ten stations and a southwesterly wind is considered.

term was added in the cost function (Eq. 21) in order to avoid oscillation of the recovered boundary conditions and in the wind stress field. This was found necessary even in the best situation when observations are available at every grid point of the model domain. Recovery of the control parameters and the boundary conditions for the first experiment but without the penalty term in the cost function is shown in Fig. 24. In both cases (Figs. 15 and 24), the pattern of recovery is similar. However, the boundary elevations do not converge to their real values when the penalty term is not added in the cost function. Furthermore, the boundary elevation (Fig. 25) oscillates in space.

The second crucial issue in the success of non-linear optimization problems is the scaling of the control variables (Section 3.6). In our study, only the inverse of the Manning's roughness is two orders of magnitude larger than the other parameters (Table 2). Therefore, scaling the inverse of the Manning's roughness by a factor 10^{-2} would seem appropriate. Experiments 1 and 2 were repeated when the mentioned scaling was adopted. While there is no difference in the estimate of the control variables when data are available everywhere, an effect of the scaling is noticeable when data are only available at ten stations. The recovered values for the wind stress reach two significant figures in the case of no scaling compared to one significant figure in the case of scaling. While the boundary elevations do not converge toward their true values, the recovery is, however, better without scaling. The surface elevation reaches 85% of the real value without scaling and 75% with scaling. The pattern of recovery with scaling (Fig. 26) is slightly different from the pattern without scaling (Fig. 18). In the case of scaling, the bottom drag coefficient parameters are adjusted at the same time as the wind stress components. Finally, the cost function (Fig. 27) continuously decreases until recovery of the wind stress components and drag coefficient. Same decreasing pattern is found for the norm of the gradient of the cost function. Overall, the full recovery of all the parameters

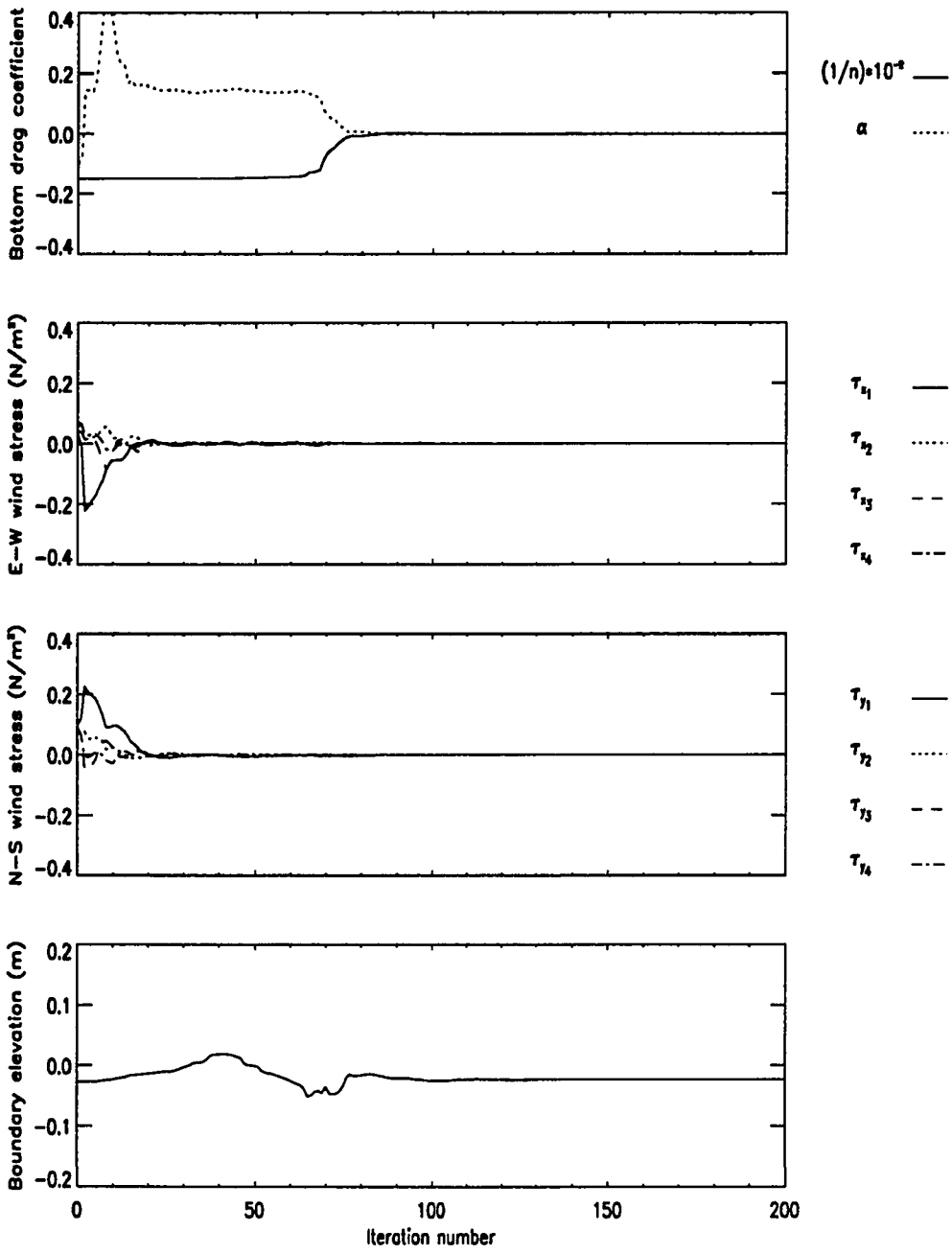

Figure 24: Difference between recovered and true values for the model parameters and boundary elevation at the southern end of the Bay mouth 24 hours after the beginning of the recovery day versus the number of minimization iterations. The data are available everywhere, a northeasterly wind is considered, and the penalty term is not added to the cost function.

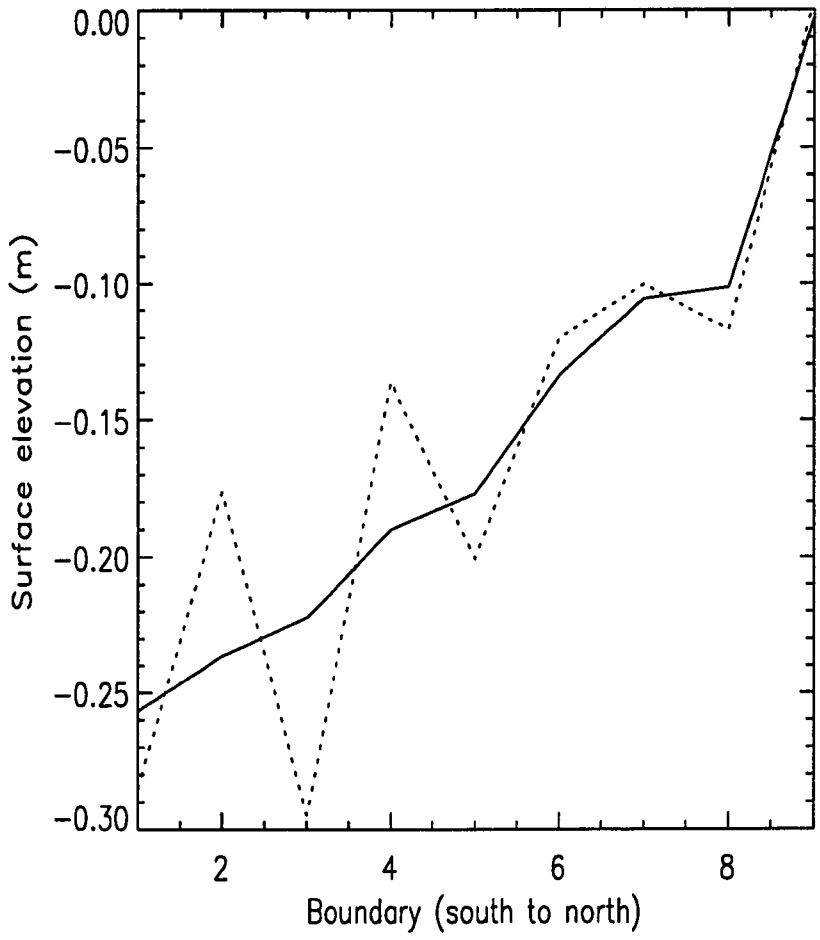

Figure 25: Recovered boundary elevation from the southern to the northern end of the Bay mouth with a penalty term in the cost function (solid line) and without the penalty term (dotted line).

is not faster but the recovered values are not as close to their true value when the inverse of the Manning's roughness is scaled.

4.2 Tidal Circulation

The success of recovery of control parameters in the identical twin experiments leads us to consider real observations in the Chesapeake Bay. As a first experiment, tidal circulation is investigated and the bottom drag coefficient parameters (Eq. 5) are estimated during the assimilation process. The assimilated observations are predicted surface elevations at several tide gauge stations in the Bay. Observations and results of the assimilation are presented in the next subsections.

4.2.1 Observations and simulation

Tidal circulation is investigated by assimilating hourly time series of sea surface elevation for a period extending from November 1 to November 19, 1983 when the most comprehensive set of simultaneous observations is available (Section 2). The hourly elevation data, referred to as predicted sea surface elevation, are predicted on the basis of five major tidal constituents, M_2, S_2, N_2, K_1, O_1, using the harmonic constants determined by Fisher (1986) (Table 3). From the deployed tide gauges in 1983, only ten tide gauges (Fig. 12) are permanent, that is, have been in place consistently for more than 30 years. Therefore, only the predicted sea surface elevations at those ten stations are used in the assimilation process and nine supplementary tide gauge stations, referred to as comparison stations, are used to compare the modeled surface elevation with the predicted observations. Locations of the tide gauges are shown in Fig. 12.

Two typical time series of predicted surface elevation at two permanent stations in the lower Bay (Chesapeake Bay Bridge Tunnel, CBBT) and in the upper Bay (Baltimore) are plotted in Fig. 28 for November 1 through November 20, 1983. Two

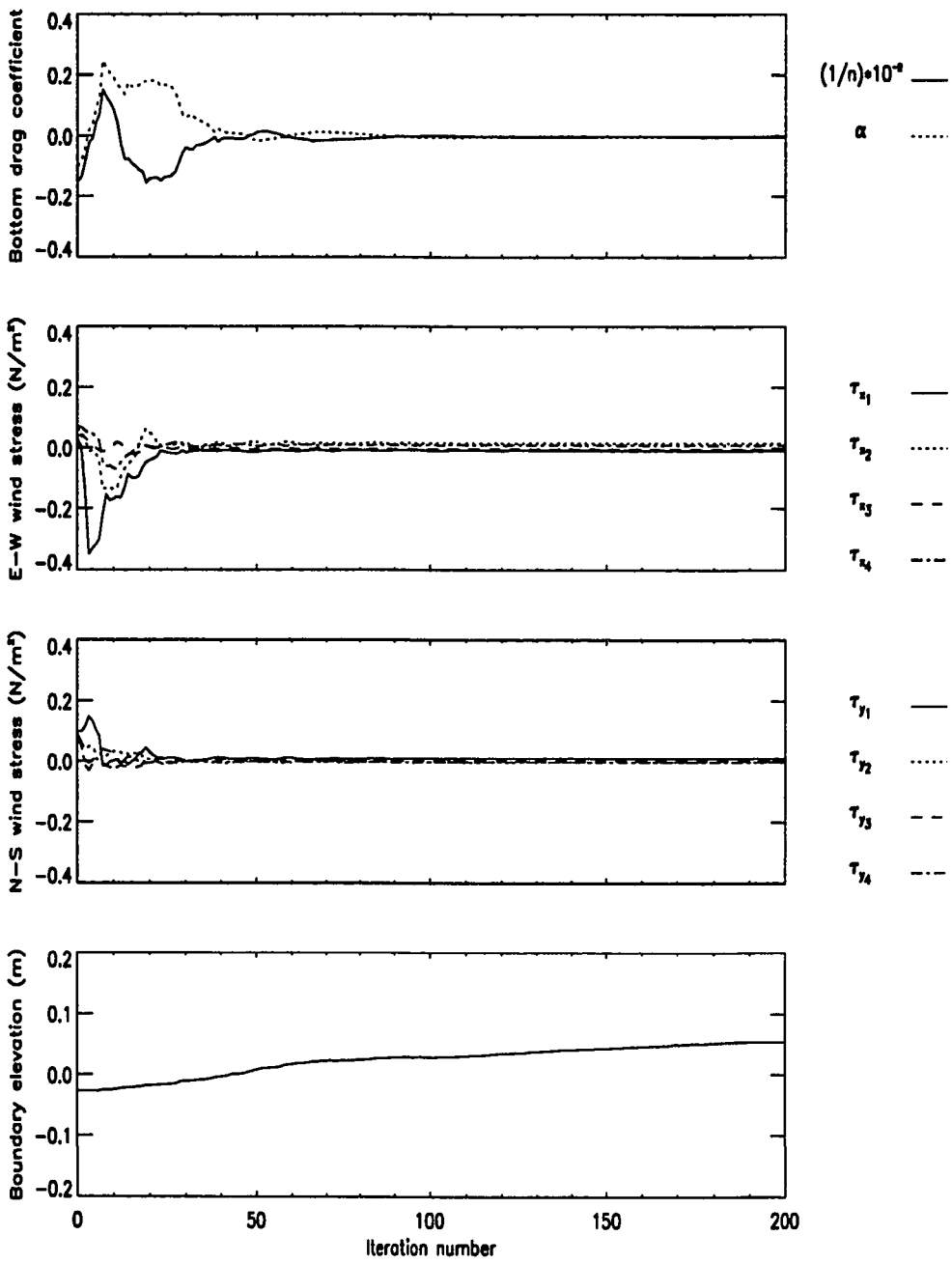

Figure 26: Difference between recovered with scaling of the Manning's roughness and true values for the model parameters and boundary elevation at the southern end of the Bay mouth 24 hours after the beginning of the recovery day versus the number of minimization iterations. The data are available at ten stations and a northeasterly wind is considered.

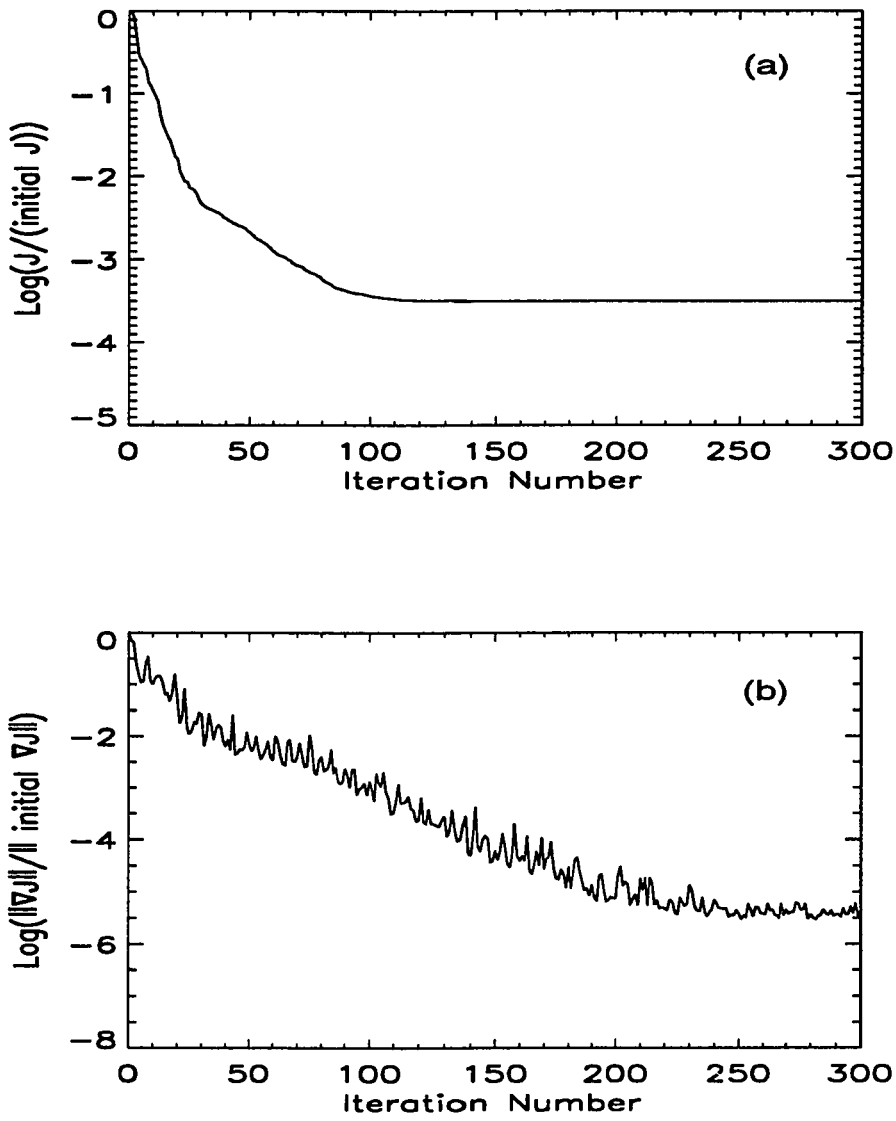

Figure 27: Logarithm of the normalized cost function (a) and normalized norm of the gradient of the cost function (b) versus the number of iterations when a northeasterly wind is blowing and the Manning's roughness is scaled.

Station Name	M_2		S_2		N_2		K_1		O_1	
	a	ϕ	a	ϕ	a	ϕ	a	ϕ	a	ϕ
Havre de Grace, MD	0.259	278.74	0.045	320.09	0.051	250.12	0.084	310.85	0.057	313.08
Baltimore, MD	0.150	186.70	0.023	216.59	0.031	162.56	0.070	294.54	0.036	294.25
Annapolis, MD	0.127	139.17	0.021	172.04	0.025	115.25	0.055	282.00	0.047	287.30
Cambridge, MD	0.231	107.45	0.032	136.84	0.044	84.06	0.052	268.54	0.045	263.76
Solomons Is, MD	0.167	45.99	0.026	70.52	0.036	19.06	0.026	240.33	0.023	256.08
Lewisetta, VA	0.179	24.94	0.030	52.72	0.036	358.96	0.023	196.67	0.020	228.09
Gloucester Pt, VA	0.347	260.14	0.063	286.78	0.078	241.26	0.048	121.87	0.033	141.41
Kiptopeake, VA	0.383	241.18	0.069	266.58	0.088	218.20	0.058	117.06	0.044	135.60
Hampton Roads, VA	0.358	255.80	0.066	285.04	0.082	236.20	0.051	124.62	0.039	143.67
CBBT, VA	0.377	229.12	0.070	254.93	0.091	208.20	0.056	110.66	0.046	130.25
Betterton, MD	0.223	253.49	0.031	292.65	0.044	221.41	0.076	301.71	0.068	312.49
Matapeake, MD	0.141	139.00	0.023	166.36	0.031	115.92	0.059	275.01	0.047	282.14
Avalon, MD	0.202	95.33	0.038	106.63	0.036	58.24	0.046	266.84	0.034	263.23
Chesapeake Bch, MD	0.137	95.66	0.022	125.37	0.030	71.44	0.044	272.45	0.038	275.71
Colonial Bch, VA	0.242	60.05	0.039	93.98	0.049	40.03	0.030	217.28	0.026	234.62
Holland Bar Lt, MD	0.203	8.41	0.036	26.49	0.042	353.70	0.027	195.88	0.029	203.74
Guardshore, VA	0.336	351.36	0.057	27.42	0.080	340.39	0.048	198.75	0.029	180.95
Rappahannock, VA	0.218	304.61	0.037	329.82	0.055	280.03	0.032	155.38	0.029	168.37
New Pt Comf Sh, VA	0.309	247.35	0.044	279.13	0.064	219.67	0.049	121.86	0.038	146.73

Table 3: Amplitude (m) and phase (deg.) for the major harmonic constituents at ten permanent and nine comparison tide gauge stations (Fisher, 1986).

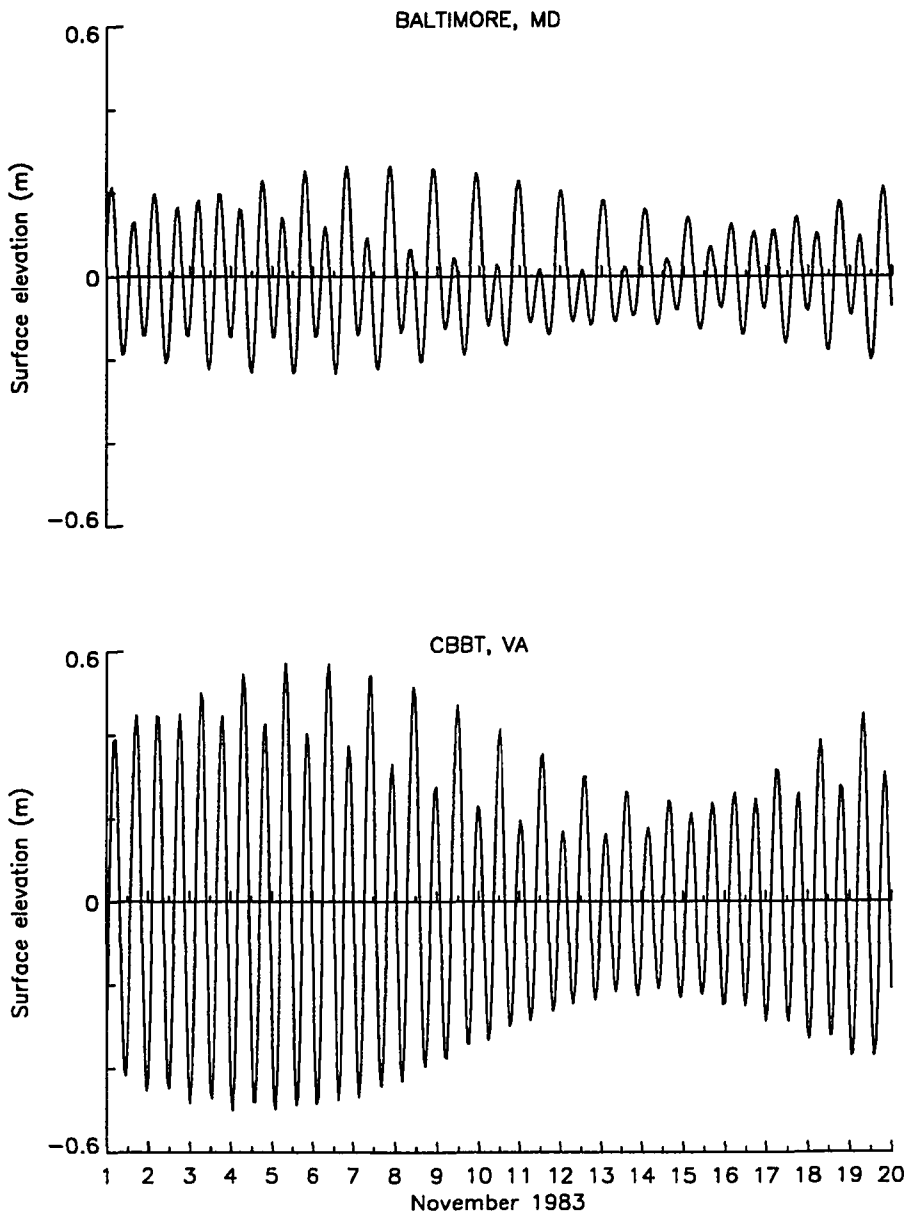

Figure 28: Time series of predicted sea surface elevation (m) at Baltimore and CBBT for November 1 to November 20, 1983.

main features can be pointed out. Spring tide occurs around November 5 when the high tide reaches its maximum while neap tide occurs around November 15 when the high tide reaches its minimum. At CBBT, the tide is mainly semi-diurnal while in Baltimore the tide is mixed mainly semi-diurnal, which is typical of the lower and upper Bay (Fisher, 1986).

A first trial of tidal circulation modeling was undertaken using the MU circulation model (Eqs. 1, 2, 3). The wind stress was set equal to zero and the quadratic bottom stress was considered with a typical constant drag coefficient of 0.002 which would be an average value for the Chesapeake Bay (Section 3.2). The circulation in the Bay was forced at the Bay mouth by imposing the predicted surface elevation based on the harmonic constants at the Chesapeake Bay Bridge Tunnel. Modeled and predicted surface elevation (Fig. 29) show poor agreement, especially in the upper Bay where modeled low and high tide amplitudes are smaller than the corresponding predicted tide. A comparison between modeled M_2 coamplitude and cophase charts (Figs. 30 and 31) and the corresponding charts from Fisher (1986) (Fig. 3) indicates that the discrepancy between modeled and observed tidal amplitudes increases towards the upper Bay. Furthermore, the modeled M_2 tide propagates slightly faster than the observed M_2 tide in the entire Bay. As a remedy to the discrepancy, one could think to decrease the drag coefficient. Decreasing the drag coefficient would indeed increase the amplitude of the tide and also change the phase lag. The question is then how much can the drag coefficient be decreased to match phase and amplitude of the M_2 tide everywhere in the Bay. One should also realize that the adjustment of the drag coefficient to match the M_2 tide would not necessarily lead to a match of the amplitude and phase of the other four constituents (Crean *et al.*, 1988). The tuning of the drag coefficient can quickly become very tedious.

An analysis of the relative average error (Eq. 23) computed over 24 hours for

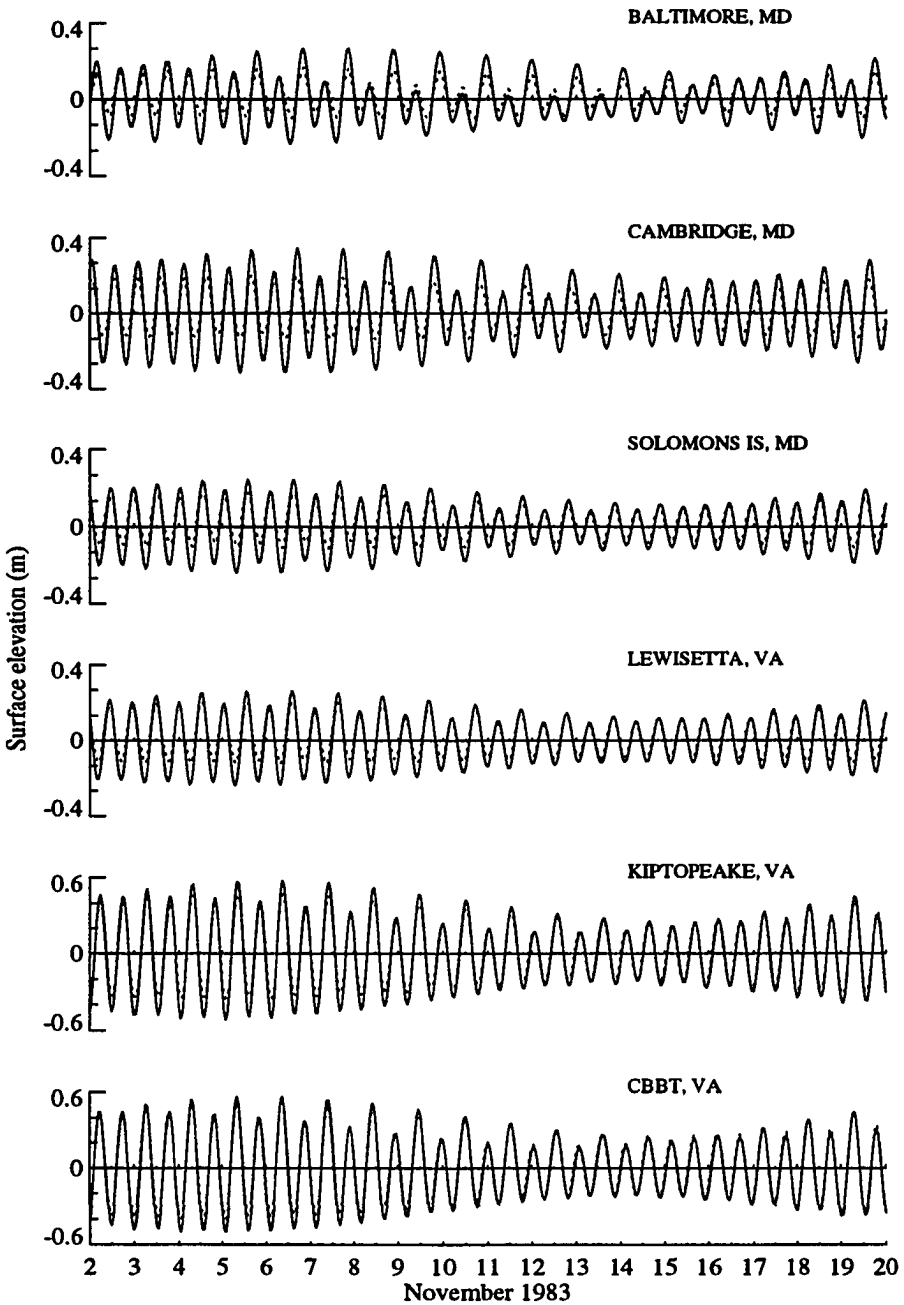

Figure 29: Time series of modeled (dotted line) and predicted (solid line) surface elevation (m) at six representative permanent tide gauge stations. The drag coefficient was taken as $c_D = 0.002$. Note the change of scale for the last two stations.

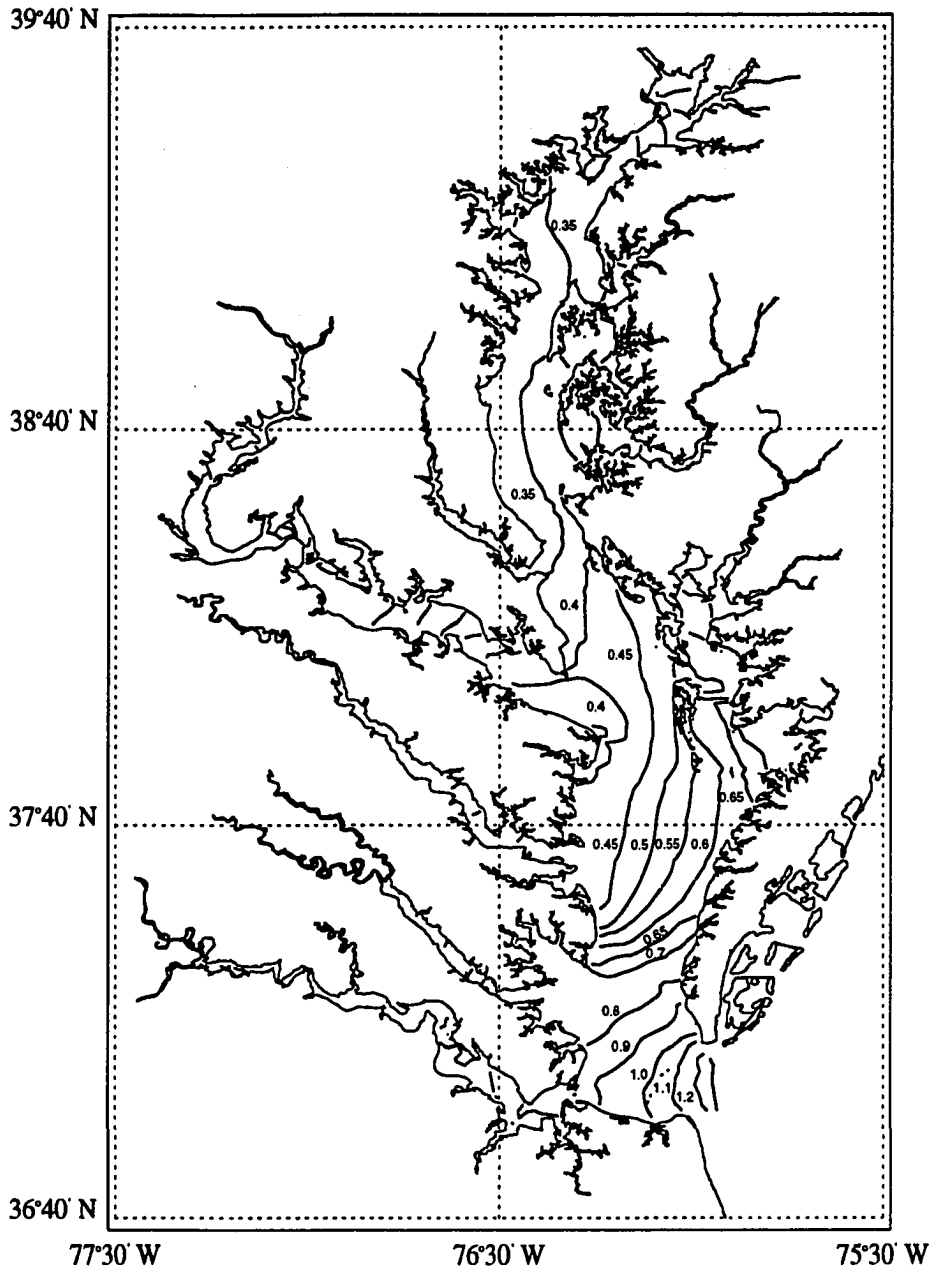

Figure 30: Coamplitude lines of the modeled M_2 tide expressed in feet.

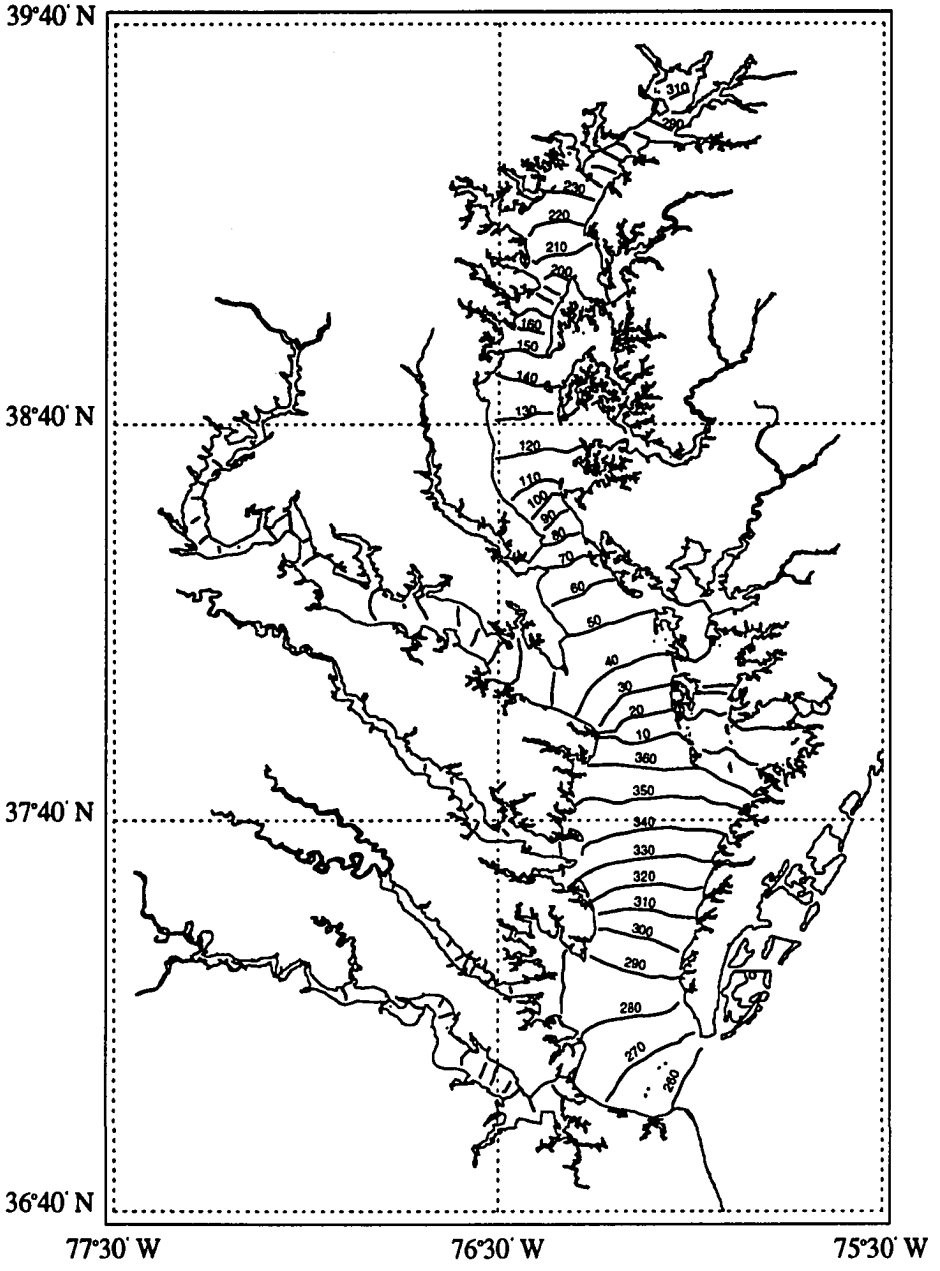

Figure 31: Cophase lines of the modeled M_2 tide expressed in degrees.

November 1 to November 19 (Table 4 and Appendix C) between modeled and predicted elevation shows an increase from a few percents in the lower Bay to about 15% in the upper Bay with a maximum in the Potomac river where the maximum relative error reaches a value of 24.3% at Colonial Beach. This variation in the relative average error from the lower Bay to the upper Bay would suggest a spatial dependency of the drag coefficient. A strong fluctuation of the relative average in time is also noted. For example, the error at Baltimore drops from a maximum of 20.7% on November 6 to a minimum of 10.7% on November 17.

Time and spatial dependence of the bottom drag coefficient was then taken into account in the following recovery experiment. A drag coefficient defined by Eq. 5 was chosen and the parameters $1/n$ and α were optimally estimated using the variational technique described previously. While those parameters are constant, the bottom drag coefficient is a function of space through its dependence in H, the total water depth. The time variation of the drag coefficient was introduced by evaluating the parameters for a period of one day for eighteen consecutive days which includes one spring and one neap tide. The actual recovery started on November 2, 1983 through November 19, 1983 while the recovery for November 1, 1983 was only used to initialize the procedure. Once the bottom drag coefficient parameters were estimated, the direct model was run for 24 hours with the new parameters in order to initialize the circulation for the following day and compare modeled and predicted elevations.

4.2.2 Recovery

For each day of the recovery experiment, the initial guesses for the parameters $1/n$ and α were taken as their estimated values from the previous day. The assimilation process was stopped when the normalized norm of the gradient of the cost function reached the value of 10^{-6}. This convergence criterion has been determined from the

results of November 1 and corresponded to the minimum value that the normalized norm of the gradient of the cost function could attain. For the period under study, convergence occurred after 9 to 18 iterations and the cost function decreased an order of magnitude. This small decrease in the cost function is due to the fact that the initial cost function was already small since the recovery was started with the best initial guess.

Predicted and modeled tidal elevations for six typical permanent and six typical comparison tide gauge stations are shown in Figs. 32 and 33, respectively. A summary of the root-mean square error, relative average error and correlation coefficient computed using Eqs. 22, 23, 24 is given in Table 4 while values for each day between November 2 and November 19 are given in Appendix C (Tables 6 to 11). Modeled elevations with the estimated drag coefficients show an excellent agreement with the predicted elevations not only at the permanent stations but also at the comparison stations. In general, the estimated low tide amplitudes are slightly smaller than the predicted amplitudes in the lower Bay. On the other hand, the recovered high tide amplitudes are slightly higher than the predicted amplitudes in the upper Bay. However, differences between modeled and predicted elevation are of the same order of magnitude as the measurement errors for the entire Bay. A maximum of less than 7% for the relative average error (Table 4) is found in the main stem while the relative error in the Potomac river (Colonial Beach) reaches a maximum of 9.86%. Moreover, the variation between spring tide and neap tide in the relative error that was found with a constant uniform drag coefficient is greatly reduced (Figs. 34, 35 and Appendix C). A correlation coefficient larger than 0.96 is found in the main stem and a minimum correlation coefficient equal to 0.91 is found in the Potomac river. This indicates a very small shift between modeled and predicted elevations, which could result from the fact that modeled elevations have been linearly interpolated in space to the location of the tide gauge. Finally, it should also be pointed

	November 2 to November 19, 1983								
Station Name	rms error (cm)			relative error (%)			correlation		
	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Havre de Grace, MD	2.4	11.8	7.31	1.73	16.90	9.70	0.955	0.990	0.973
	2.8	9.1	6.48	2.16	6.15	4.73	0.958	0.986	0.970
Baltimore, MD	3.3	7.6	5.48	10.66	20.74	15.89	0.936	0.977	0.965
	2.1	4.6	3.11	1.67	6.47	3.92	0.951	0.989	0.975
Annapolis, MD	1.7	6.4	4.13	4.54	17.73	11.33	0.969	0.985	0.975
	1.1	4.6	2.65	0.93	7.03	3.43	0.979	0.996	0.989
Cambridge, MD	4.5	9.8	7.16	8.94	17.16	13.78	0.945	0.996	0.974
	3.4	4.0	3.80	1.85	6.71	3.53	0.956	0.995	0.979
Solomons Is, MD	2.5	7.4	4.76	7.01	16.30	10.97	0.961	0.988	0.976
	2.0	4.9	3.25	2.09	5.38	4.07	0.962	0.992	0.983
Lewisetta, MD	1.9	6.7	4.10	3.30	11.32	6.76	0.985	0.997	0.993
	1.2	5.1	3.03	0.96	4.55	2.72	0.988	0.998	0.993
Gloucester Pt, VA	2.6	12.5	7.29	1.79	10.44	5.87	0.971	0.997	0.985
	1.6	7.3	4.31	0.62	2.92	1.73	0.989	0.999	0.995
Kiptopeake, VA	3.4	8.4	5.26	0.97	3.41	2.32	0.976	0.999	0.993
	2.9	4.1	3.47	0.54	2.51	1.11	0.976	1.000	0.993
Hampton Roads, VA	2.9	11.0	6.51	2.01	7.36	4.26	0.990	0.998	0.994
	2.7	7.4	4.73	1.44	2.92	2.05	0.988	0.999	0.996
CBBT, VA	2.3	5.8	3.72	0.40	1.85	1.17	0.983	0.999	0.995
	2.9	4.4	3.49	0.60	2.25	1.09	0.980	1.000	0.994
Betterton, MD	2.4	10.9	6.97	2.78	19.65	11.67	0.970	0.996	0.982
	2.1	6.4	4.37	1.72	4.04	2.85	0.983	0.994	0.990
Matapeake, MD	1.9	6.4	4.13	3.86	14.15	9.03	0.962	0.995	0.978
	1.5	4.3	2.72	1.71	4.57	2.76	0.986	0.996	0.993
Avalon, MD	3.2	8.6	5.48	1.34	14.34	8.56	0.955	0.994	0.983
	2.1	5.4	3.54	1.36	4.95	2.95	0.964	0.998	0.985
Chesapeake Bch, MD	1.6	5.5	3.46	2.80	11.45	6.99	0.971	0.997	0.985
	1.6	4.9	3.13	2.35	6.09	3.99	0.984	0.998	0.993
Colonial Bch, VA	3.5	12.2	7.77	6.59	24.29	15.58	0.924	0.976	0.950
	2.9	9.6	6.08	3.77	9.86	6.72	0.910	0.968	0.939
Holland Bar Lt, MD	2.4	6.8	4.31	3.09	8.81	5.70	0.966	0.995	0.985
	2.0	4.7	3.22	1.72	3.24	2.52	0.974	0.998	0.991
Guardshore, VA	3.8	10.5	6.70	3.27	8.06	5.32	0.967	0.992	0.982
	3.0	6.6	4.77	1.41	3.14	2.20	0.973	0.991	0.983
Rappahannock, VA	3.0	6.2	4.15	2.50	5.57	4.39	0.950	0.995	0.978
	3.1	4.3	3.61	1.83	5.07	3.07	0.958	0.998	0.983
New Pt Comf Sh, VA	2.6	8.2	5.20	1.38	5.36	3.49	0.979	0.997	0.991
	2.3	5.2	3.67	1.03	2.07	1.58	0.988	0.998	0.994

Table 4: Minimum, maximum and mean value of the root-mean square error (cm), the relative average error (%) and the correlation coefficient for November 2 to November 19, 1983. The first ten stations are the permanent stations while the last nine are the comparison stations.

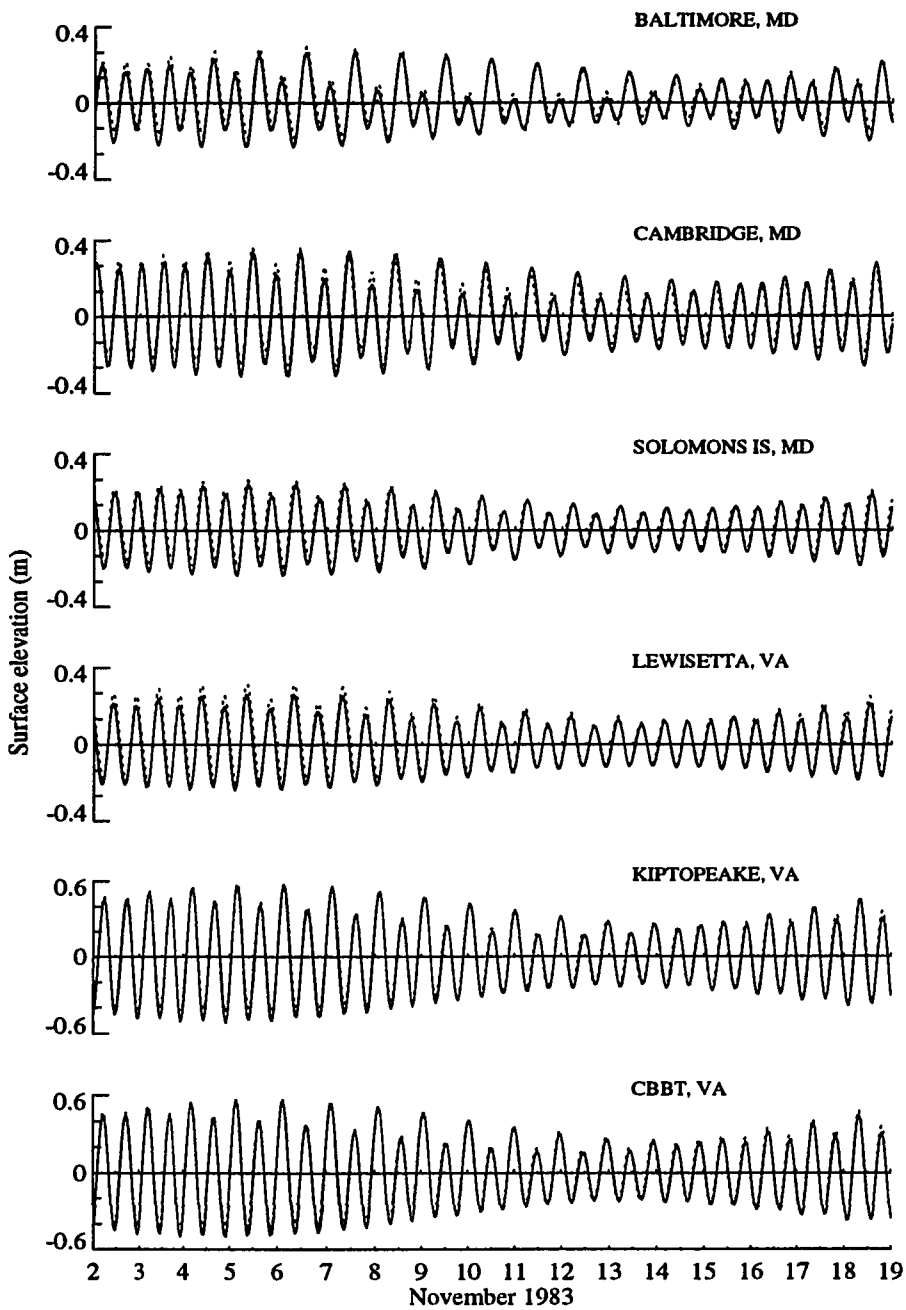

Figure 32: Time series of modeled (dotted line) and predicted (solid line) surface elevation (m) at six permanent tide gauge stations. Note the change of scale for the last two stations.

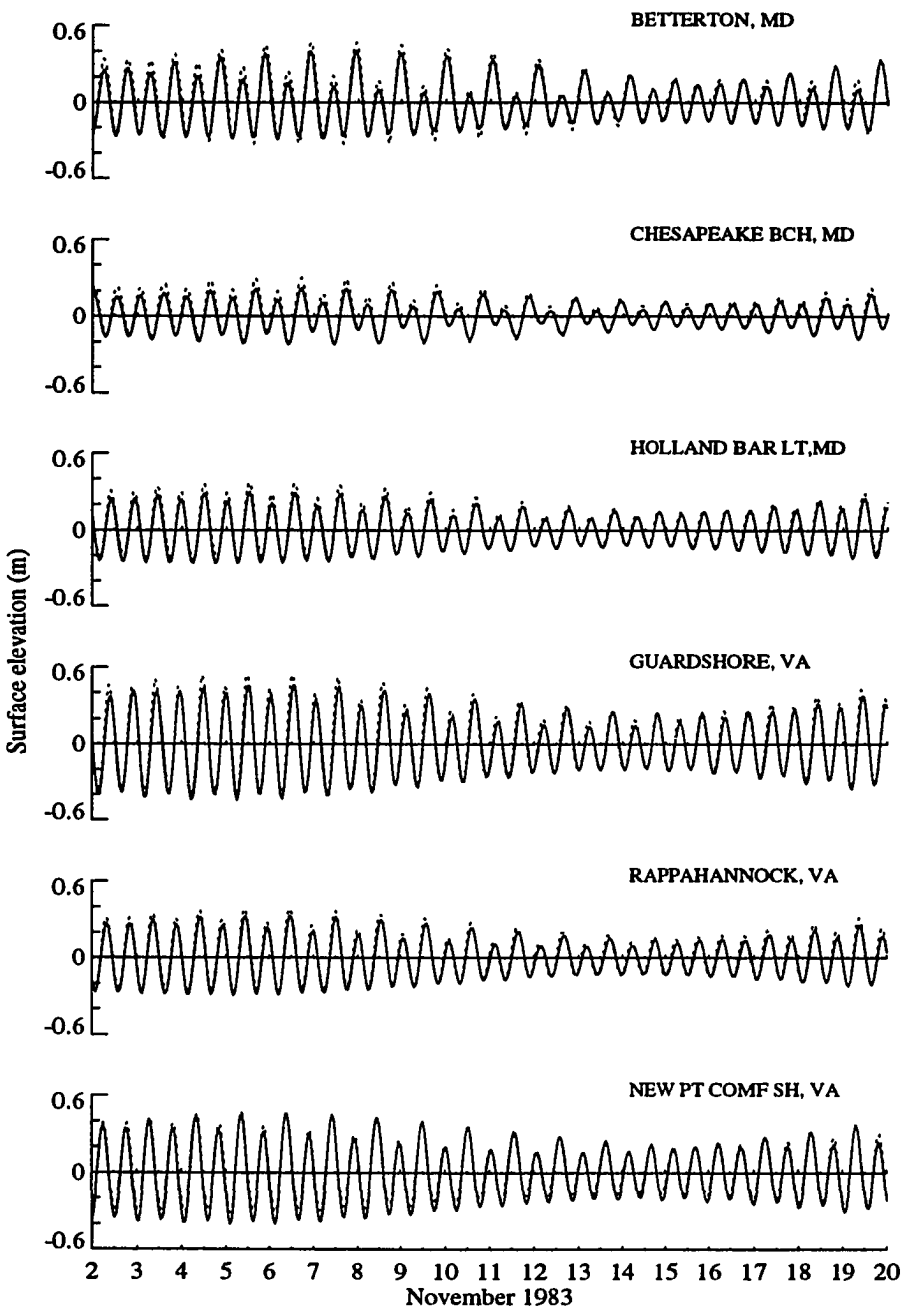

Figure 33: Time series of modeled (dotted line) and predicted (solid line) surface elevation (m) at six comparison tide gauge stations.

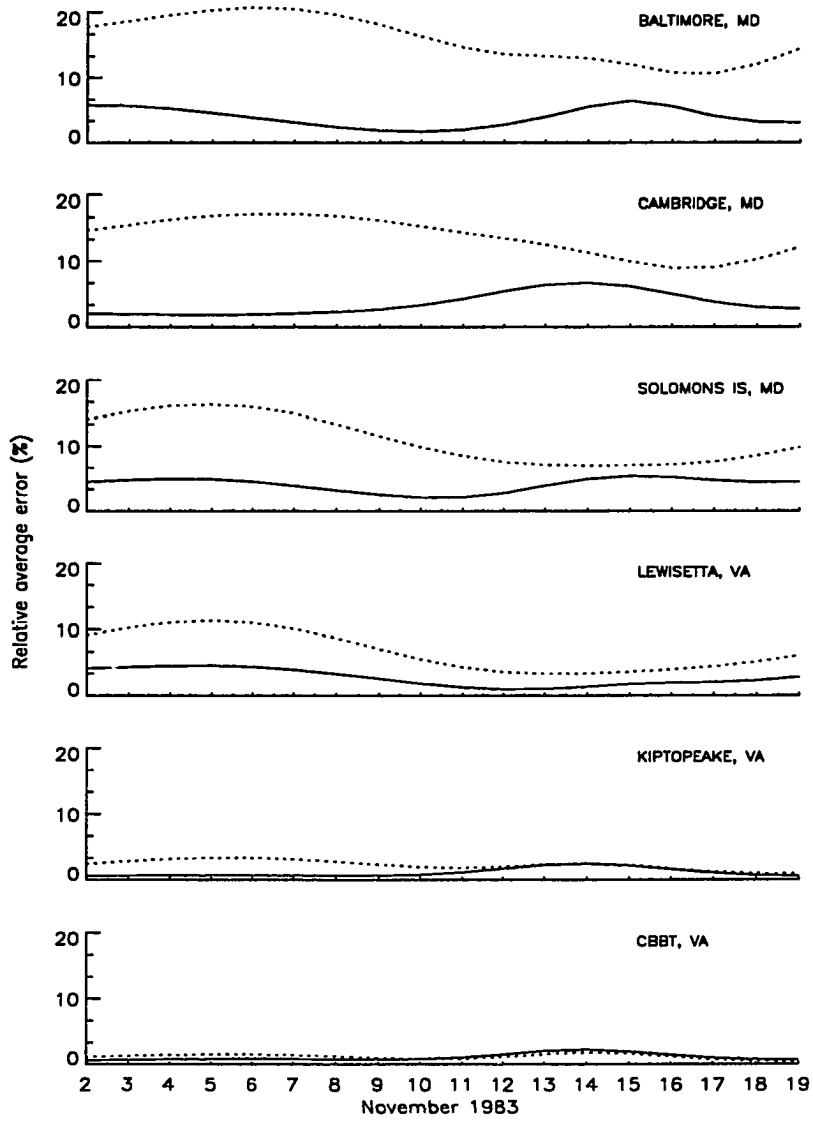

Figure 34: Time series of relative average error (%) for recovered (solid line) and modeled ($c_D = 0.002$) (dotted line) surface elevations at six permanent tide gauge stations.

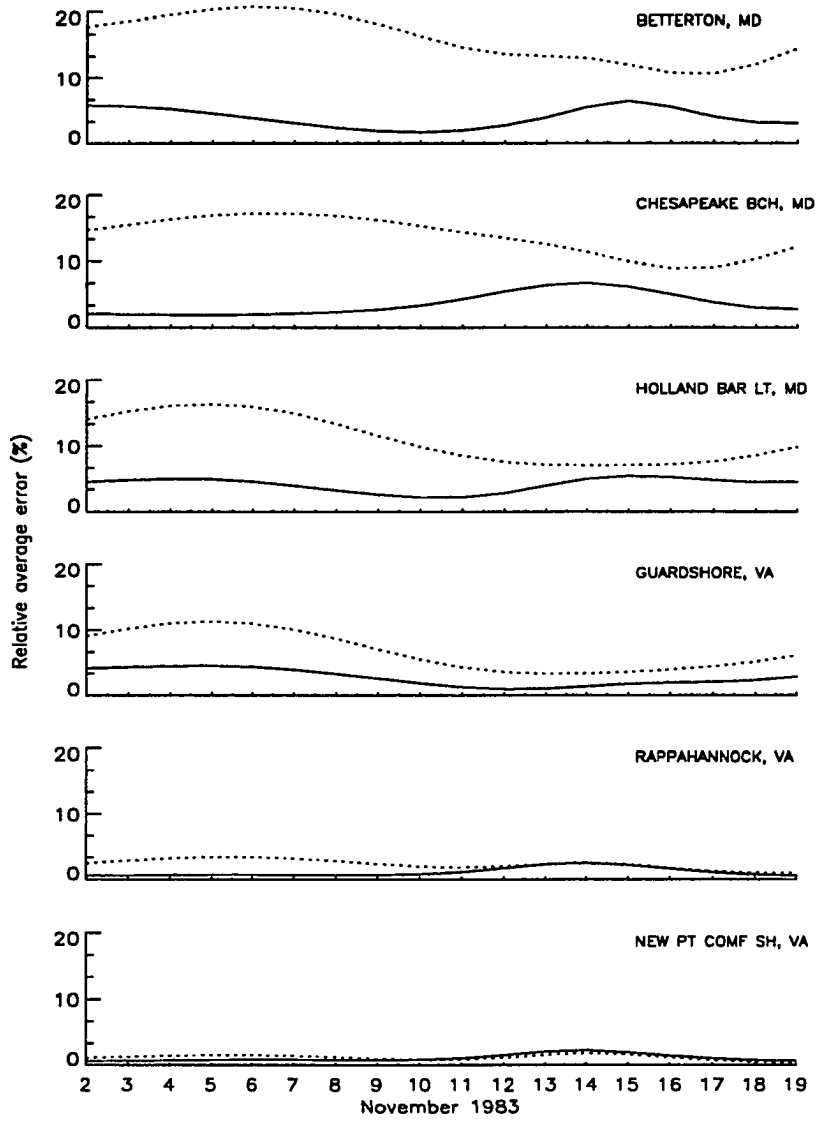

Figure 35: Time series of relative average error (%) for recovered (solid line) and modeled ($c_D = 0.002$) (dotted line) surface elevations at six comparison tide gauge stations.

out that the same agreement is found at Havre de Grace despite the fact that the station is near the Susquehanna river which has the biggest discharge of all rivers.

Using harmonic analysis for the time series of the modeled elevations with the estimated drag coefficients between November 2 and November 19, the amplitude and phase of the M_2 and K_1 tide at every grid point of the model domain were computed. A comparison between the M_2 coamplitude chart from Fisher (1986) (Fig. 3) and the corresponding chart for the modeled M_2 (Fig. 36) shows a very good agreement. Same agreement (not shown) was found for the K_1 tide. The M_2 coamplitude lines run across the Bay at the entrance, then become longitudinal and finally run across the Bay in the upper Bay. The M_2 amplitude decreases from the Bay mouth towards the main stem, and then increases north of Baltimore. The amplitude of modeled M_2 tide is slightly larger than the amplitude found by Fisher (1986), which can be attributed to the fact that the time series of modeled elevations (18 days) was too short to fully separate the M_2, N_2 and S_2 tides. A comparison between the M_2 cophase chart from Fisher (1986) (Fig. 3) and modeled M_2 cophase chart (Fig. 37) again shows excellent agreement. For example, the 310 degree cophase line goes through Rappahannock in both cases. The cophase lines are uniformly spaced in the lower Bay. In the upper Bay, the cophase line spacing decreases in the narrow portions and increases in the wider portions. The fact that the curvature of the recovered cophase lines at the entrance of the Bay is not as pronounced as the curvature found by Fisher (1986) is due to the imposed boundary condition at the Bay mouth. The predicted elevation was taken uniform at the Bay mouth which does not allow any curvature of the cophase lines. Two virtual amphidromic points, *i.e.*, near the Severn river (north of Annapolis, Fig. 12) and the Potomac river, are evident. At those locations, the coamplitude lines are concentric and the cophase lines converge. Those two points were also found by Browne and Fisher (1988).

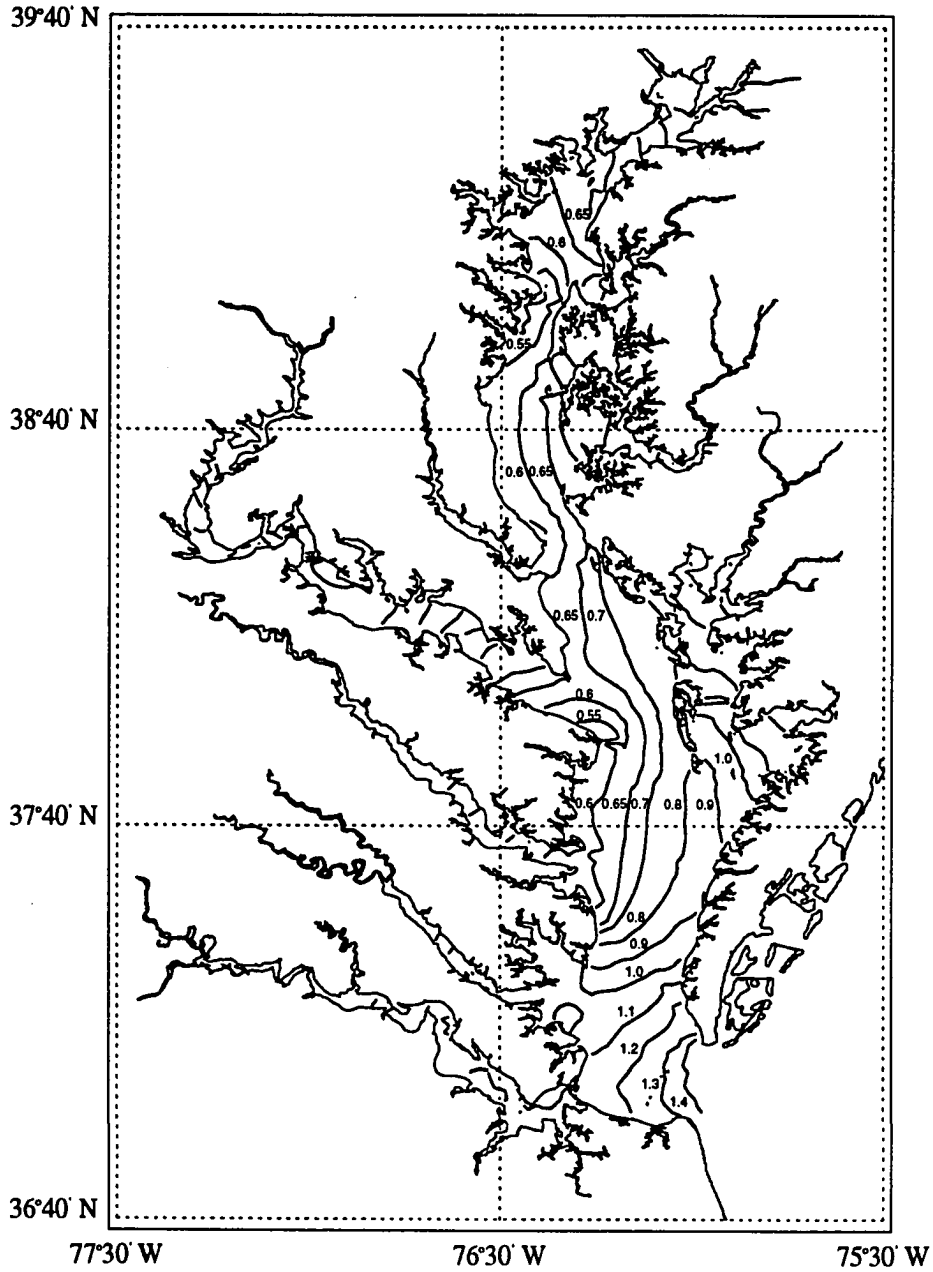

Figure 36: Coamplitude of the recovered M_2 tide expressed in feet.

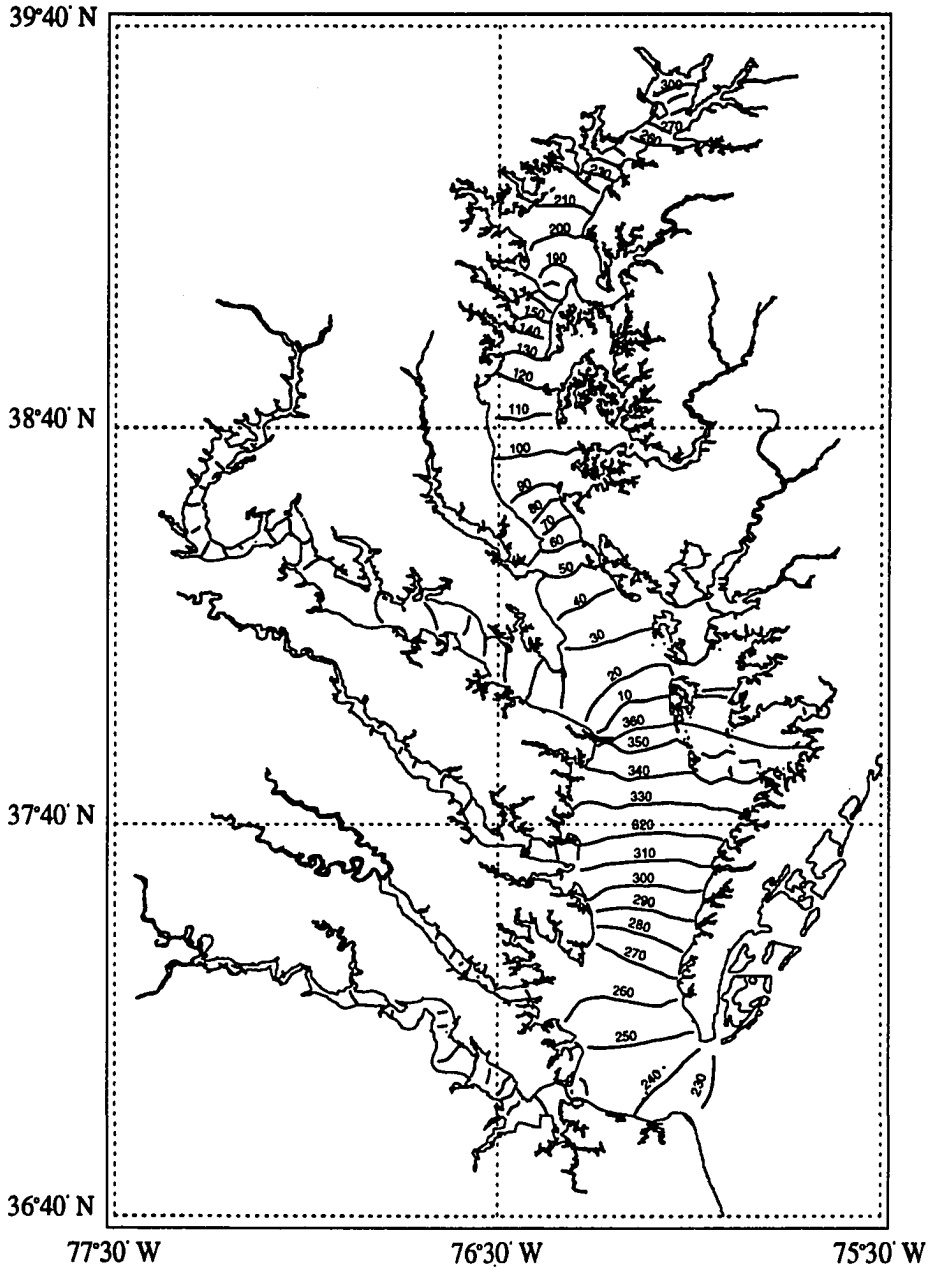

Figure 37: Cophase of the recovered M_2 tide expressed in degrees.

While predicted and modeled elevations are in excellent agreement, probably the most interesting feature to note is the periodicity of the estimated inverse Manning's roughness and therefore of the bottom drag coefficient. As shown in Fig. 38, the bottom drag coefficient decreases during spring tide (day 5-6) and increases during neap time (day 14-15). The exponent α (Eq. 5) is mainly constant during the entire period.

The difference between spring and neap tide values is larger in the regions with a depth less than 10 m. The estimated bottom drag coefficient varies between 2.5×10^{-4} and 3.1×10^{-3}. The lower value is smaller than the values found in the classical tidal modeling literature (Ronday, 1976; Werner and Lynch, 1987; Crean *et al.*, 1988). However, using an inverse method and tidal-current observations in the lower and upper Bay, Bang (1994) found a similar range of values for the bottom drag coefficient, *i.e.*, $2.0 \times 10^{-4} \leq c_D \leq 1.6 \times 10^{-3}$. Discussion of the eventual causes of the periodicity of the drag coefficient can be found in Section 5.

4.3 Wind-driven circulation

Since modeled tidal circulation in the Bay has shown to be considerably improved by using variational data assimilation, our next focus is on wind-driven circulation which is even harder to model. Indeed, the surface forcing, *i.e.*, the wind stress, is poorly known for the Chesapeake Bay since the wind observations are available only at the major airports and do not necessarily represent the wind over the water. As mentioned in Section 3.6, the goal of the present experiment is to estimate the bottom drag coefficient, the wind stress components at four locations and the hourly surface elevations at the open boundary during the assimilation process and get the best representation of the wind-driven circulation in the Bay. As for the tidal circulation study, the assimilation period is in November. However, in order to have wind observations to compare to the estimated wind using the variational

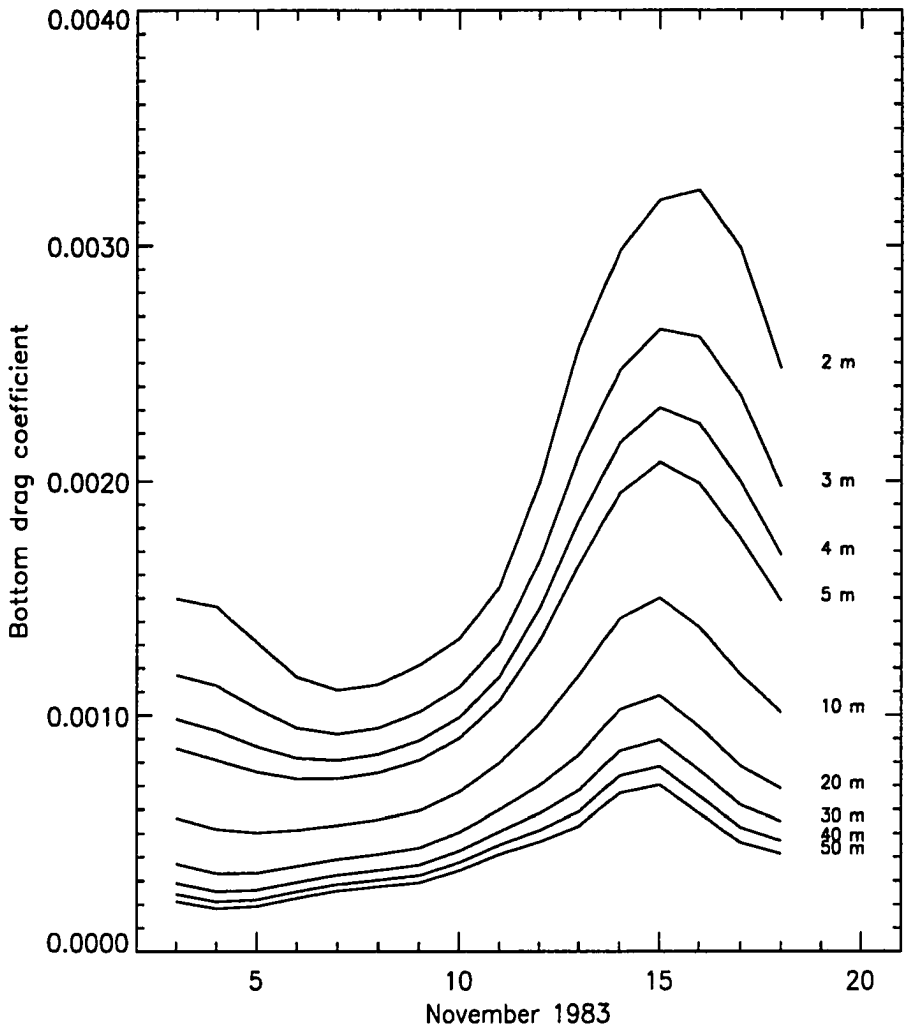

Figure 38: Time series of estimated bottom drag coefficient c_D for depths between 2 and 50 m.

assimilation technique, the year 1990 had to be considered instead of the year 1983. Observations and results of the assimilation process are described in the next two subsections.

As additional evidence of the improvement of the modeled wind-driven circulation by using data assimilation, two days in September 1983 when data are available at more tide gauges than the one used for assimilation were also considered. Hence, it was possible to compare modeled elevations with observations which were not assimilated. Unfortunately, during that period of time, wind speed and direction was only measured at the airports. The results of that experiment is presented in the last subsection.

4.3.1 Observations

Wind-driven circulation in the Chesapeake Bay was investigated by assimilating hourly sea surface elevation observations from ten permanent tide gauges (Fig. 12) between November 2 and November 8, 1990. As in the tidal circulation experiment, observations of November 1 are used to initialize the recovery process. During that period of time, hourly wind speed and direction observations were available at two buoys deployed by NOAA (Fig. 12), Thomas Point and Chesapeake Light Tower, and at the tide gauge station CBBT. A comparison between observed and estimated wind was therefore possible.

Time series of predicted using the harmonic constants from Fisher (1986) and observed sea surface elevations at two permanent tide gauge stations, Baltimore and CBBT, are plotted in Fig. 39 for November 1 through November 10, 1990. Differences between predicted tidal and observed elevations are noticeable during those few days, which suggests that the wind plays an important role. During the chosen assimilation period, either a southwesterly or southeasterly wind (Fig. 40) persisted for about four days over the entire Bay with a speed less than 5 m/s in

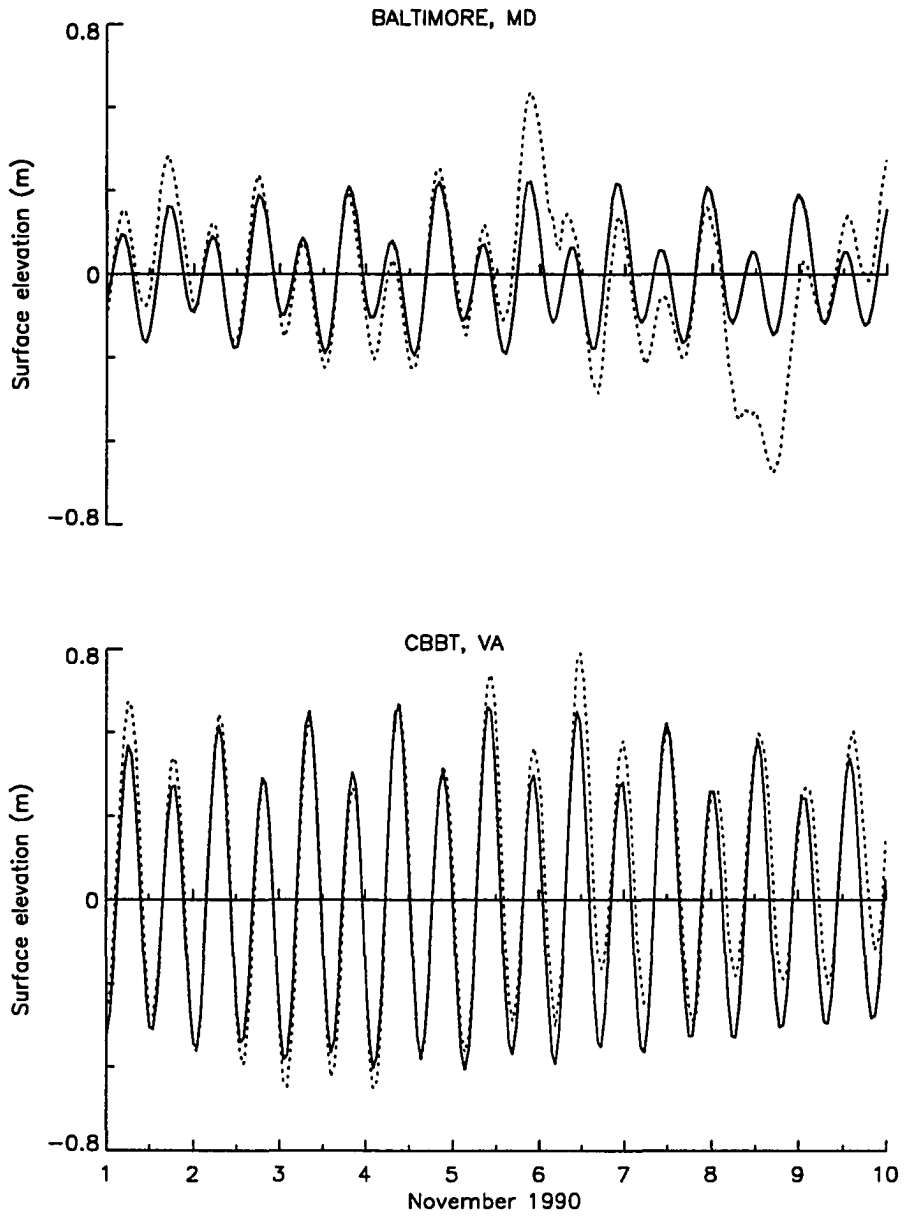

Figure 39: Time series of predicted tidal (solid line) and observed (dotted line) elevation in Baltimore and CBBT for November 1 to November 10, 1990.

the upper Bay and slightly higher in the lower Bay. During those few days, a sea breeze pattern is also noticeable. In the afternoon of November 5, the wind speed increased to about 10 m/s in the upper Bay. An increase of the surface elevation is immediately noticeable in the upper Bay, *e.g.*, the high tide elevation is about twice its predicted value at Baltimore. A smaller change in the surface elevation can also be seen at CBBT. On November 6, a change of the wind direction associated with the passage of a cold front, is evident. The wind became northwesterly in the upper Bay and northeasterly in the lower Bay. Again, a similar effect of the wind direction change can be seen in the observed surface elevation at both stations. The surface elevation decreased in Baltimore and increased at CBBT, which is the expected response to a northwesterly and a northeasterly wind, respectively. This change in the surface elevation is larger on November 8 when the wind become stronger. Finally, notice that the wind direction outside of the Bay (Chesapeake Light Tower) is roughly the same at CBBT while the wind speed is slightly larger. Therefore, we do not anticipate any different behavior in the surface elevation from the Bay mouth and at CBBT.

The time variations of the wind stress and the bottom drag coefficient were introduced in the recovery experiment by evaluating the parameters for a period of one day for seven consecutive days. The actual recovery started on November 2 through November 8, 1990 while the estimated parameters for November 1, 1990 were only used to initialize the procedure. Once the bottom drag coefficient parameters, wind stress and boundary conditions were estimated, the direct model was run for 24 hours with the new parameters in order to initialize the circulation for the following day and compare modeled and predicted elevations.

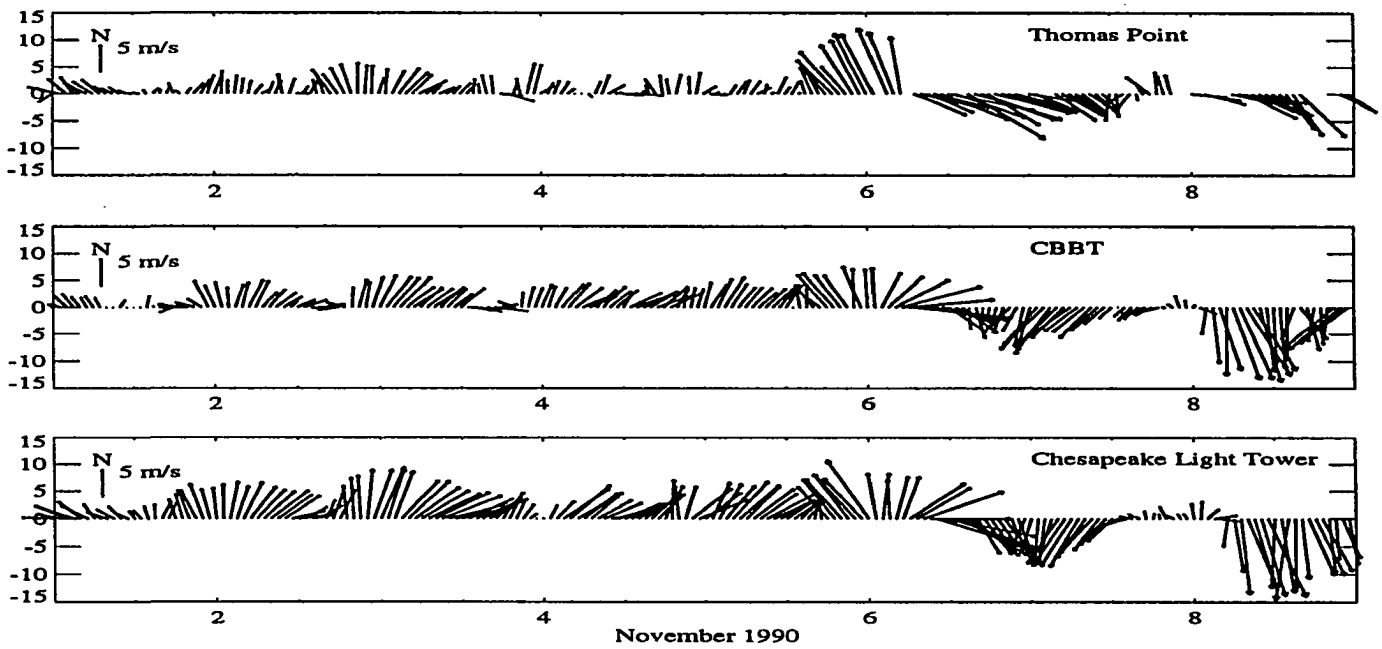

Figure 40: Hourly observed wind in November 1990 at two buoys, Thomas Point and Chesapeake Light Tower, and at the tide gauge station CBBT. The stick diagram is plotted using the oceanographic convention.

4.3.2 Recovery

For each day of the assimilation experiment, the initial guess for the control variables was taken as their estimated value from the previous day. Based on the recovery on November 1, the recovery was stopped when the normalized norm of the gradient of the cost function reached a value of 10^{-3}. The convergence criterion was set to a higher value than for the tidal experiments. The assimilated data are indeed actual observations and are noisier than the data used in the tidal recovery. For each day of assimilation, the cost function decreased one order of magnitude. As for the tidal experiment, the small decrease in the cost function can be attributed to the fact that the assimilation runs were started with the best initial guess. The number of iterations necessary to satisfy the preset convergence criterion varied between 244 and 370. The cost function decreased rapidly during the first 30 iterations and continued to slowly decrease until convergence was reached. This pattern was also found for the twin experiments (Section 4.1).

Time series of modeled sea surface elevation is shown in Fig. 41. The relative average error over 24 hours at six permanent tide gauge locations is plotted in Fig. 42. Root-mean square error, relative average error and correlation coefficient (Eqs. 22, 23, 24) are given in Table 5 for every day of the assimilation period. Excellent agreement between estimated and observed surface elevation is evident for the entire period of assimilation in the lower Bay. In general, a relative average of error less than 3% is found for the stations in the lower Bay and the correlation coefficient is over 0.99. In the upper Bay, the agreement is also very good until the frontal passage. The relative average error is less than 5% until November 6 when it increases to about 15%. While the correlation coefficient is over 0.9, a variation in its value is also noticeable for the same period.

Recovered wind vectors at Thomas point and CBBT for November 2 to November 8, 1990 are plotted in Figs. 43 and 44, respectively. Wind speed and direction

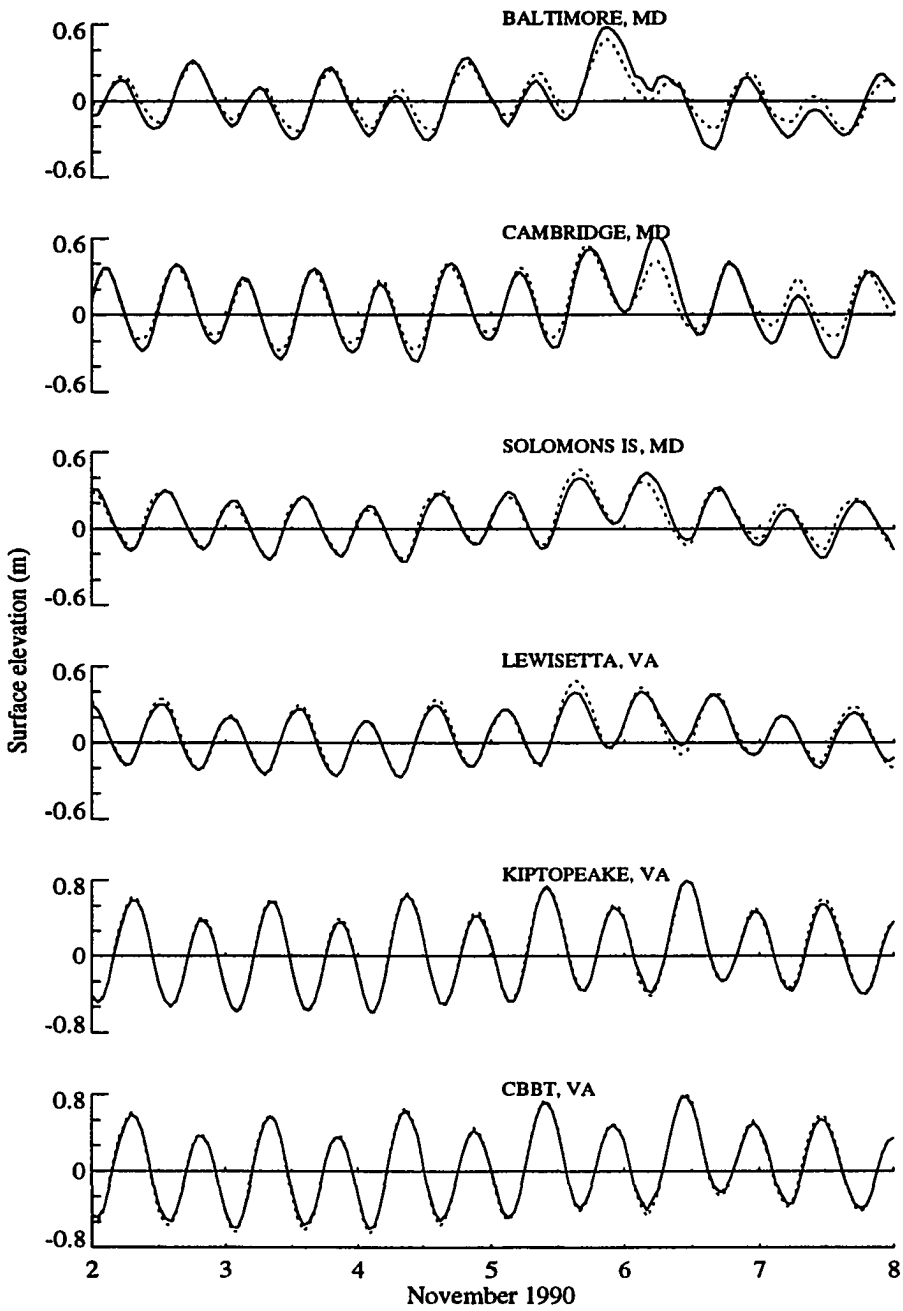

Figure 41: Time series of modeled (dotted line) and observed (solid line) surface elevation (m) at six permanent tide gauge stations.

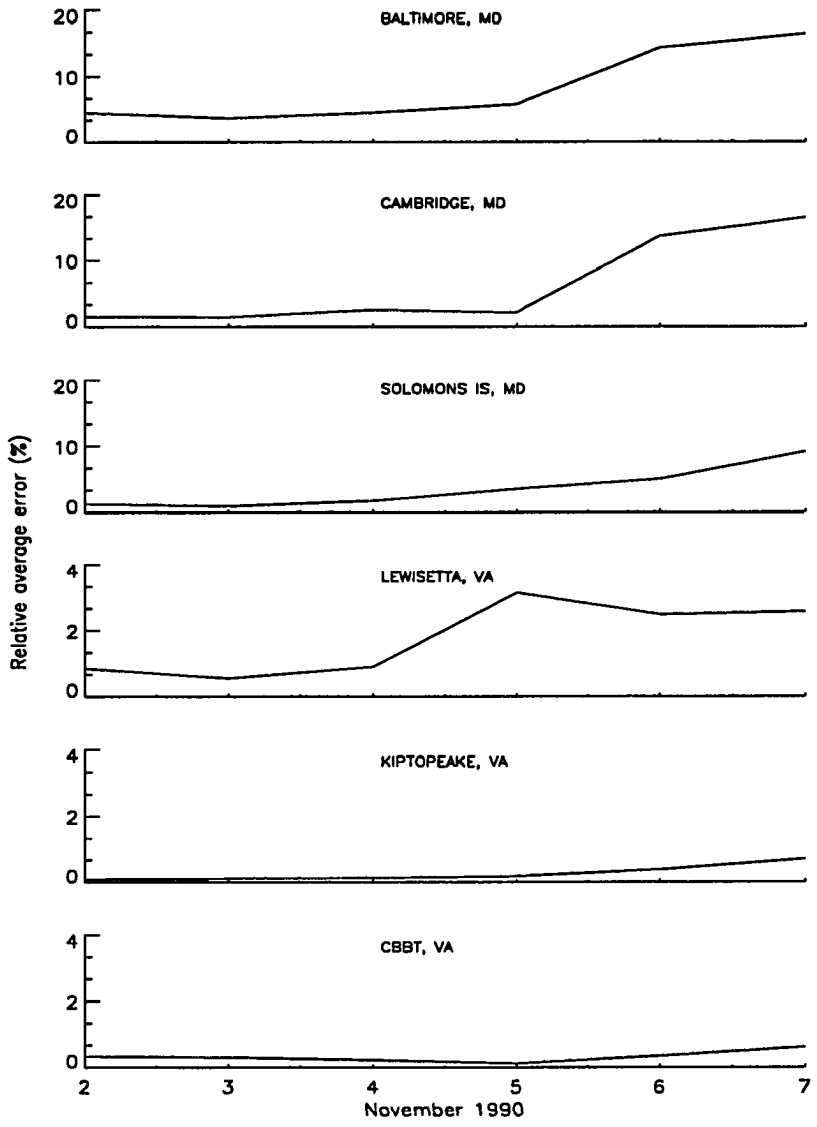

Figure 42: Time series of relative average error (%) between modeled and observed surface elevation at six permanent tide gauges. Note the change of scale for the last three stations.

Station Name	November 1990						
	1	2	3	4	5	6	7
Havre de Grace, MD	5.3	6.2	4.3	7.9	9.2	5.8	1.1
	1.73	3.03	1.24	3.44	3.71	1.77	11.51
	0.985	0.971	0.991	0.968	0.972	0.985	0.895
Baltimore, MD	4.0	4.5	4.1	5.5	7.5	9.5	8.0
	3.78	4.46	3.61	4.46	5.71	14.29	16.43
	0.976	0.983	0.990	0.986	0.976	0.937	0.897
Annapolis, MD	4.9	2.0	2.4	2.1	3.6	8.5	5.0
	7.11	1.03	1.44	0.91	1.72	14.91	7.49
	0.970	0.990	0.994	0.991	0.983	0.926	0.949
Cambridge, MD	6.0	3.8	3.7	5.0	4.8	11.2	11.2
	4.82	1.56	1.45	2.58	2.14	13.64	16.47
	0.986	0.994	0.994	0.994	0.988	0.930	0.943
Solomons Is, MD	3.9	2.6	2.2	2.9	4.4	5.4	5.8
	3.40	1.26	0.91	1.73	3.53	5.02	9.21
	0.974	0.990	0.992	0.987	0.978	0.960	0.957
Lewisetta, MD	2.2	2.3	1.9	2.4	4.5	3.8	3.3
	0.92	0.85	0.55	0.89	3.15	2.48	2.58
	0.992	0.995	0.996	0.995	0.987	0.982	0.984
Gloucester Pt, VA	4.1	4.0	3.8	3.4	2.8	6.5	5.5
	0.92	0.82	0.65	0.48	0.35	2.13	1.93
	0.992	0.994	0.994	0.996	0.997	0.985	0.982
Kiptopeake, VA	1.8	1.6	1.8	2.1	2.3	3.2	3.8
	0.13	0.09	0.11	0.13	0.18	0.38	0.71
	0.999	0.999	0.999	0.999	0.998	0.997	0.995
Hampton Roads, VA	3.5	4.1	4.2	3.6	3.0	5.3	5.6
	0.63	0.79	0.73	0.50	0.36	1.32	1.84
	0.998	0.997	0.997	0.999	0.998	0.995	0.986
CBBT, VA	2.3	3.1	3.1	2.7	1.9	3.0	3.5
	0.22	0.33	0.30	0.22	0.12	0.34	0.61
	0.999	0.998	0.998	1.000	1.000	0.999	0.996

Table 5: Root-mean square error (cm) (first number), relative average error (%) (second number), and correlation coefficient (third number) for the wind-driven experiment

were obtained from the wind stress components using a quadratic law with a drag coefficient of 2.0×10^{-3} (Schwab, 1982). The wind pattern is well represented in the lower Bay. At CBBT, the estimated wind is mainly southwesterly and shows an increase of magnitude late on November 5. A change of direction occurs on November 6, which is in agreement with the observed wind pattern. In the upper Bay, the agreement is not as good until the frontal passage when the wind speed increased to 10 m/s and a change in direction occurred. This would suggest that during periods of weak wind, mechanisms other than the local wind forcing are more important in the narrowest part of the Bay.

The recovered drag coefficient (Fig. 45) shows a minimum on November 5, which corresponds to the period of spring tide. This is followed by an increase until November 6 when there is a frontal passage. The drag coefficient then decreases and reaches roughly the same value regardless of the depth of the considered region in the Bay. While a continuous increase of the drag coefficient was expected until neap tide, northwesterly and northeasterly winds seem to decrease the magnitude of the drag coefficient compared to its value with no wind or southeasterly wind.

The last recovered control variable to examine is the boundary elevation at the Bay mouth. As expected, the high tide elevation (Fig. 46) is higher and the low tide is lower on the northern end of the Bay mouth than at the southern end. This difference in elevation decreased when the wind changed direction and blew from the northwest and the northeast. The elevation at the northern end of the Bay mouth is slightly higher than the expected value. This can be due to the fact that the boundary region in our model is about half the width of the real boundary (Fig. 9).

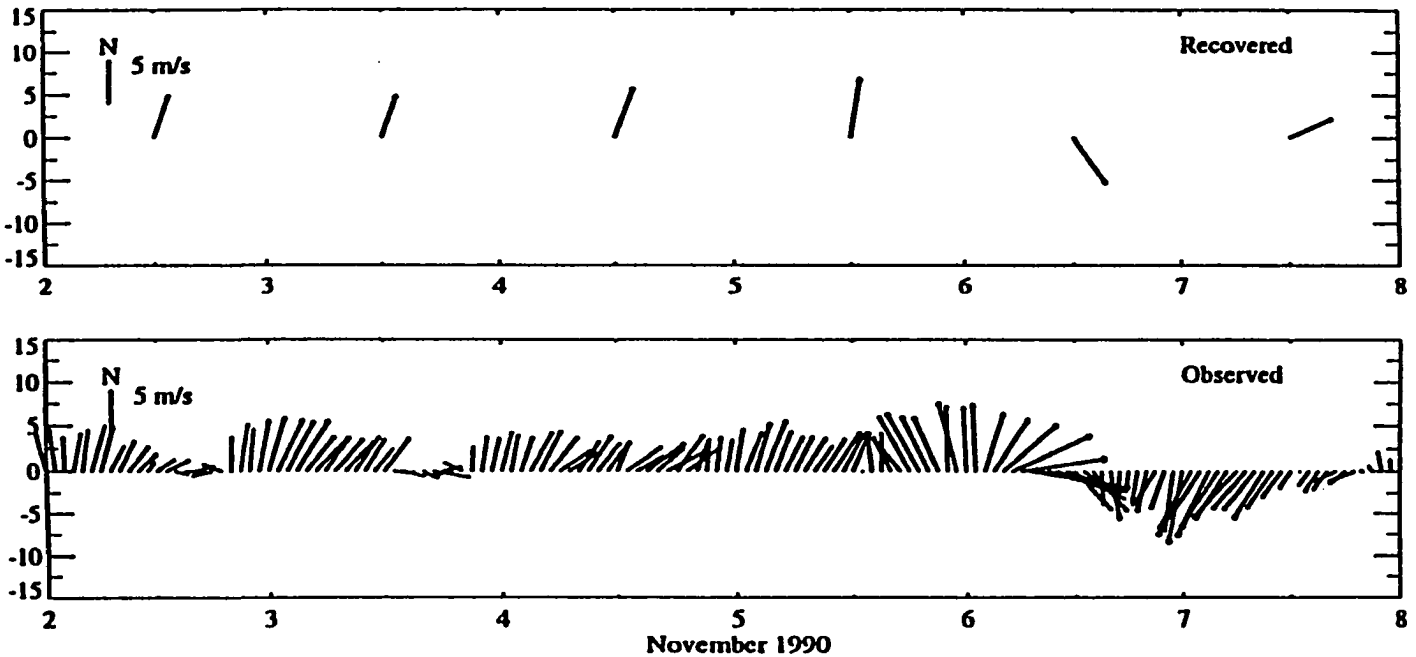

Figure 44: Recovered and observed wind at CBBT. The recovered wind, plotted at the middle of the recovery day, is constant during that day.

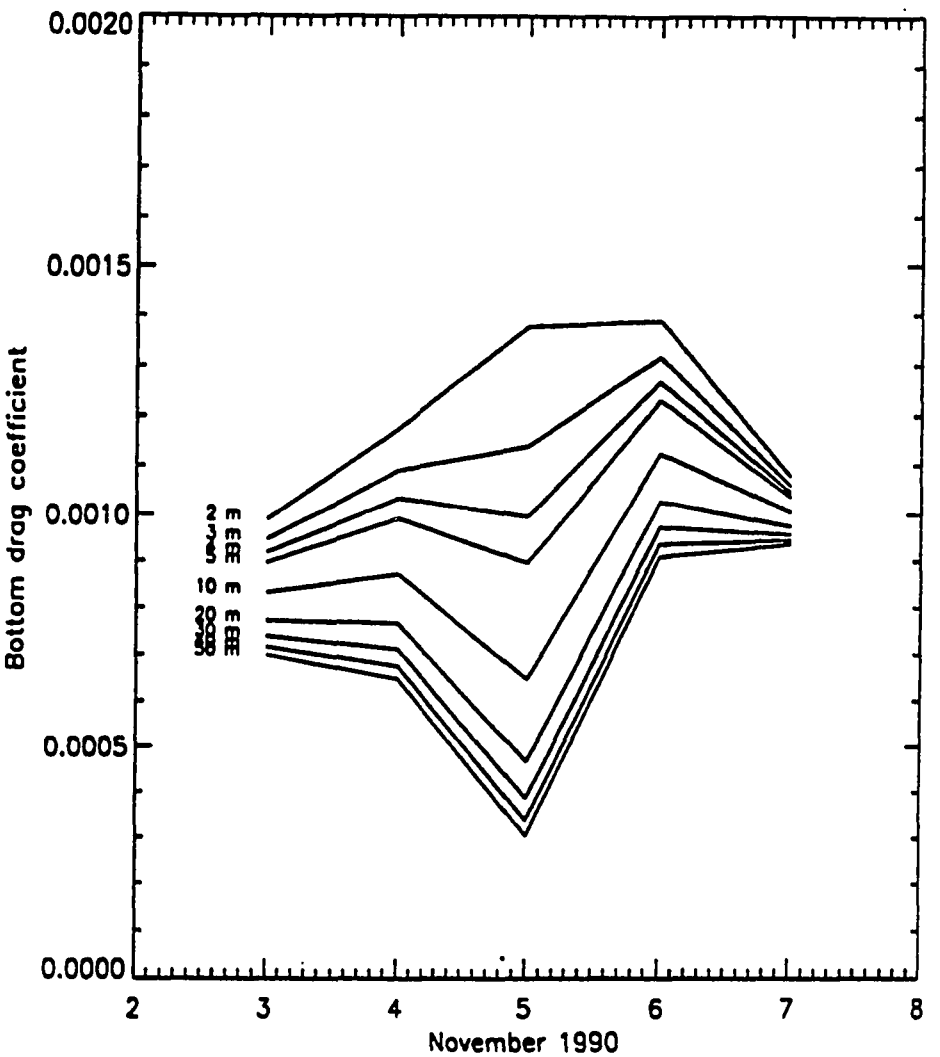

Figure 45: November 1990 time series of recovered bottom drag coefficient c_D for depths between 2 and 50 m.

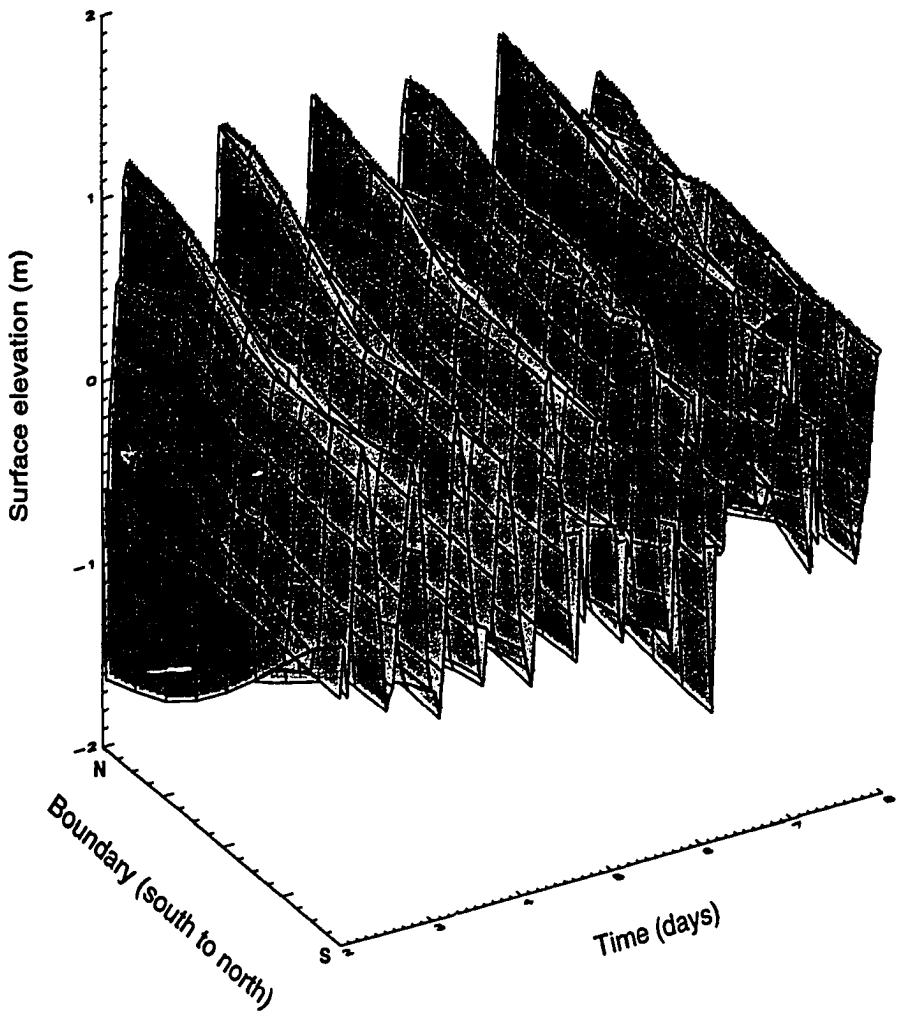

Figure 46: Time series of hourly boundary surface elevation (m) from the southern end to the northern end of the Bay mouth.

4.3.3 Further investigation of the recovery

In the previous wind-driven circulation experiment, sea level was only measured at ten tide gauges and all the observations were used in the assimilation process. In the best case, one can expect that at those ten stations there is no misfit between the modeled elevations obtained using the estimated parameters and the observations. Indeed, the assimilation procedure is based on the minimization of the misfit between modeled and observed elevations. However, this does not guarantee no misfit at other locations in the Bay. In order to increase our confidence in the feasibility of recovery of wind-driven circulation in the Bay, the previous recovery experiment has been repeated for September 20 and September 21, 1983. During that period of time, sea level measurements are also available at tide gauges other than the permanent tide gauges. It was then possible to compare modeled and observed surface elevations from other tide gauges than the one used during the assimilation process. Unfortunately, the wind observations are only available at the major airports and comparison between estimated and observed wind is very difficult. Time series of the predicted based on five major constituents and observed elevations at two tide gauge stations, Baltimore and Chesapeake Bay Bridge Tunnel, are plotted in Fig. 47. During the considered period of time, a dominant effect of the wind on the surface elevation can be seen. One should also point out that at the same time there was a destratification of the Bay (Blumberg and Goodrich, 1990). Therefore, the circulation during those two days should be well represented by our barotropic model.

The modeled sea surface elevation with the estimated bottom drag coefficient is plotted in Fig. 48. Modeled and observed surface elevations show a very good agreement not only at the permanent but also comparison tide gauges. Furthermore, the magnitude and direction of the recovered middle Bay wind, *i.e.*, southwesterly wind, was found to be comparable to the wind measured at the Patuxent River Naval

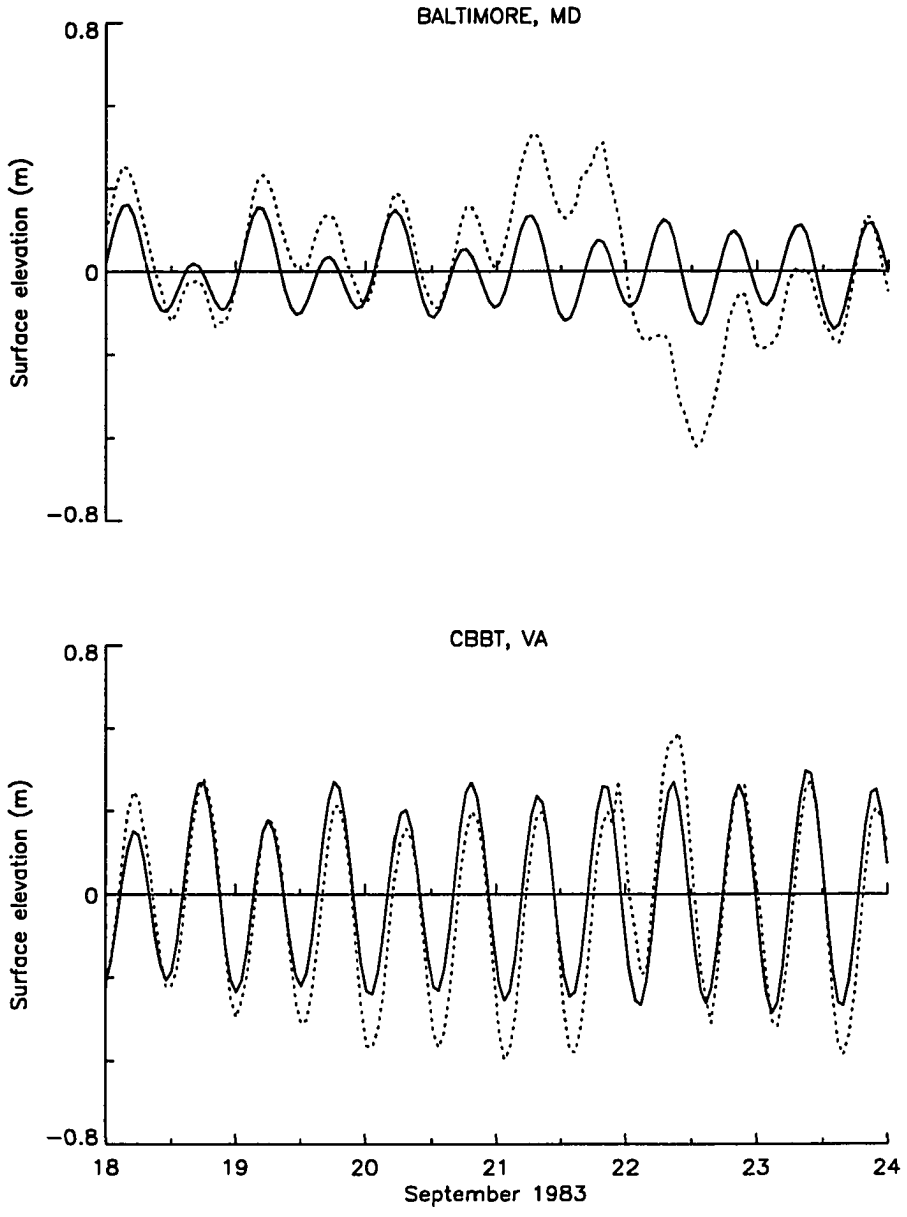

Figure 47: Time series of observed (dotted line) and predicted (solid line) elevation at Baltimore and CBBT from September 18 to September 24, 1983.

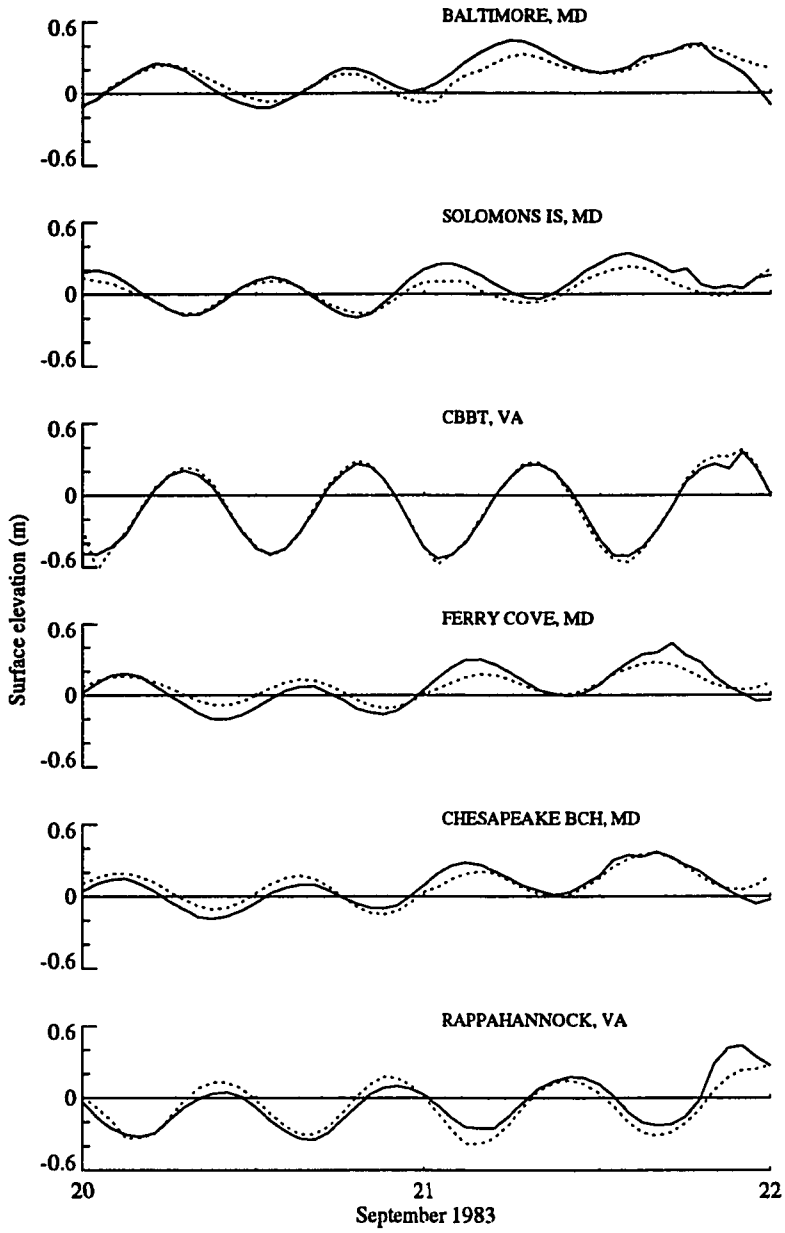

Figure 48: September 1983 time series of recovered (dotted line) and observed (solid line) surface elevation (m) at three permanent and three comparison tide gauge stations.

Air station (Blumberg and Goodrich, 1990). This experiment indicates that the modeled wind-driven circulation in the main stem can be improved by assimilating only the surface elevations measured at the ten permanent tide gauges.

5 Discussion

The previous sections presented the formalism, implementation and results of adjoint variational assimilation of tide gauge observations in the Chesapeake Bay when the dynamical constraint is a 2-D vertically-integrated shallow water model. The following sections first describe the problem of identifiability and regularization in parameter estimation. The results of the data assimilation experiments are then discussed from two perspectives: first from the point of view of the assimilation technique and second from the improvement of the modeled circulation in the Chesapeake Bay. There was no attempt to improve the assimilation technique nor the circulation model but rather to use those two tools and tide gauge observations to estimate the model control variables which could improve the modeled tidal and wind-driven circulation in the Bay.

The discussion of the identical twin experiments focuses on the definition of the cost function. This is followed by a discussion of the rate and precision of the recovery of the control parameters and boundary conditions when model generated observations are subsampled. The results of the tidal and wind-driven circulation experiments are discussed within the context of the physics of the Bay. Particular attention is given to the estimated bottom drag coefficient for tidal and wind-driven circulation and to the recovered wind stress in the wind-driven circulation. Finally, future studies are indicated.

5.1 Identifiability and regularization in parameter estimation

The main problem in parameter estimation comes from the identifiability of parameters, especially when the number of parameters is large. In general, identification refers to the determination of unknown parameters such that the predicted response of the model is close to the process observations. Banks and Kunish (1989) and Omatu and Seinfeld (1989) addressed the question of identifiability of distributed parameter systems and proposed performance criteria, *e.g.* least-squares performance criterion (cost function). Due to the ill-posedness of the problem, regularization approach is often used which leads to a well-posed problem. The problem of regularization has been defined by Omatu and Seinfeld as follows. *Regularization* of a problem refers to solving a related system called the regularized problem, the solution of which is more regular in a sense than that of the original problem, and which approximates the solution of the original problem. Regularization is thus an approach to circumvent lack of continuous dependence on the data. The regularized problem is a well-posed problem whose solution yields a physically meaningful answer to the given ill-posed parameter estimation problem (Kravaris and Seinfeld, 1985). Banks and Kunish (1989) pointed out that the addition of the regularization term changes the nature of the problem and that the solutions of the regularized problem are different from those of the original problem. Adding a regularization term to the fit-to-data criterion can, for instance, specify certain additional smoothness properties of the solution. Finally, experience with parameter estimation shows that severe difficulties arise when using unconstrained unregularized algorithms with the adjoint approach to cstimate unknown parameters when the dimension of the approximating state is kept fixed while increasing dimension of approximating parameter spaces (Kunish and White, 1986; Yeh, 1986).

In this study, the important parameters to be estimated were identified from

previous studies of the circulation in the Chesapeake Bay (Section 2). The regularization approach consists in the use of penalty-regularization terms in the cost function. The effect of the addition of a penalty term to the cost function is discussed in the following section.

5.2 Data assimilation and twin experiments

First and foremost, the identical twin experiments provide insight into the feasibility of parameter estimation in the Chesapeake Bay, in the use of an appropriate cost function, and in the adequacy of the observations that are not contaminated with observational errors.

5.2.1 Definition of the cost function

As mentioned in Section 3, variational data assimilation consists of estimating the best value of the control parameters by minimizing a cost function which measures the misfit between model results and observations. The main concern in parameter estimation using variational data assimilation is the adequacy of the type and number of observations used to define the cost function. In his discussion of linear regression as a paradigm of model fitting, Thacker (1987) showed that the minimum number of observations must be at least as large as the independent model variables. However, even if the number of observations is large enough, adequacy of the data is not guaranteed. Tziperman *et al.* (1992a) showed that even though the number of temperature and salinity observations was larger than the number of unknown wind stress components, those observations did not provide independent information and therefore the convergence to the correct forcing was relatively slow. They showed that adding surface velocity observations improved the convergence to the correct wind forcing. Similar problems can arise from the discretization on fine grids required to resolve the phenomenon under study. Indeed, the observations

might be too sparse in space and time to determine the gridded variables. Adding observations is not always possible. A cure is to supplement the inadequate data with prior knowledge or prejudice that the gridded values should be smooth. This is equivalent to adding penalty terms to the cost function, which is defined as the weighted square of the differences between model results and observations. As mentioned in Section 3.6, Richardson and Panchang (1992) and Lardner *et al.* (1993) showed that it was necessary to add a penalty term to the cost function in order to assure smoothness of the estimated field.

In this study, the identical twin experiments show that in order to recover a smooth open boundary condition close to its real value, a term, J_p (Eq. 21), which penalizes large variations of the surface elevation from one grid point to another at the boundary, need to be added to the cost function (Eq. 20). The necessity of the penalty term is evident even when there is a maximum of observations, *i.e.*, data at every grid point. In that case, the number of observations is more than ten times the number of unknown variables. This suggests that surface elevation does not provide enough information about the boundary elevation, which can be explained as follows. The exchange between the open ocean and the Bay through the Bay mouth has a strong influence on the surface elevation in a limited area of the lower Bay even though the signal generated at the boundary propagates throughout the entire Bay. Furthermore, the elevations at the Bay mouth have a smaller phase lag with the elevation in the lower Bay than with the elevation in the upper Bay. Therefore, the intensity of the signal from the boundary conditions will be stronger in the surface elevation observations in the lower Bay than in the upper Bay. This explanation is also supported by the distribution of the relative average error between observations and recovered elevation after 15 iterations of the assimilation process (Figs. 17, 20, 23). After 15 iterations, when wind stress components almost reach their true values, the effect of the boundary condition is

clearly seen in the lower Bay where the relative average error is larger than elsewhere in the Bay.

Instead of penalizing the cost function, one could increase the window of assimilation to two days, for example, which would result in an increase in the number of observations. However, a larger window would not solve the problem since the number of required boundary conditions will also increase. One could assimilate different types of observations, *e.g.*, current velocities. This is not possible in the Chesapeake Bay since the current is not routinely measured. The addition of the penalty term thus seems the most appropriate for our study.

5.2.2 Scaling of the control variables

The second concern in the variational method comes from the use of non-linear optimization, such as the limited memory quasi-Newton method. The control variables often differ by several orders of magnitude and the scaling of those variables has been shown to be important in order to accelerate the convergence. Discussion of the role of various scaling techniques in improving the performance of descent algorithms can be found in Gill *et al.* (1981), Navon and de Villiers (1983). Seiler (1993) showed that it is important to have the right scaling of the control variables otherwise the optimization might fail to converge, indeed yield senseless parameters. In this study, scaling similar to the one used by Navon *et al.* (1992a) was applied. This scaling (Section 3.6) brings all the control variables to the same order of magnitude. The only unknown parameter, two orders of magnitude larger than the other parameters, is the inverse of the Manning's roughness. Dividing it by a factor 100 is the simplest way of bringing all the parameters to the same order of magnitude (Section 3.6). However, when scaling of the inverse of the Manning's roughness parameter is applied, the true value of the parameters is not recovered. On the other hand, good results in an acceptable number of iterations was obtained

without scaling of the parameters. In that case, wind stress components are first adjusted followed by the drag coefficient parameters and finally the boundary elevations. For this study, scaling of the largest parameter does not help the recovery process and therefore is not used. Investigation of different scalings of the control variables was beyond the scope of the study. But, that problem needs to be addressed in a future study, especially when the number of parameters is larger than in the cases treated here.

5.2.3 Rate and precision of the recovery

The rate and precision at which the control variables are recovered in the twin experiments clearly show a dependence on the availability of observations. Over the past several decades, only ten permanent tide gauges (Section 2 and Fig. 12) have been deployed in the Chesapeake Bay. Most of those tide gauges are situated in embayments or entrances to rivers. It is, however, very encouraging to see that even with limited observations, *i.e.*, with hourly surface elevation observations at ten stations, it is possible to recover, to two significant figures, the bottom drag coefficient parameters, the wind stress components at four locations in the main stem while boundary conditions are recovered to 85% of their true values. For all the twin experiments, the wind stress components are first recovered, which indicates that the wind stress is the dominant forcing in the circulation model. As previously mentioned, the lack of recovery of the boundary conditions to their true values can be due to the fact that the signal from the boundary is strong only in the surface elevation measured at the tide gauges in the lower Bay. In addition, the number of tide gauges near the Bay mouth (Fig. 12) is rather small. Indeed, only two stations, CBBT and Kiptopeake are in the main stem while two other stations, Hampton Roads and Gloucester, are located ear the Bay mouth in the entrance to rivers. The surface elevation at the latter two stations would also reflect

the effect of the river discharge, of the geometry of the rivers and of bottom and surface stresses, which are probably more important than the effect of the boundary conditions. A way of increasing the importance of the boundary elevation signal during the recovery would be to weight differently the data misfit between modeled and observed elevations from the lower Bay Proper stations in the cost function. In this study, there was no attempt to give more weight to those observations to compute the cost function which could lead to a better recovery of the boundary condition. The boundary is indeed arbitrarily taken some distance from the Bay mouth and is also narrower than the real opening at that distance. Therefore, we are not interested in the recovery of the boundary condition itself but in an estimate of the boundary condition which would be physically acceptable and lead to the best circulation in the Bay.

5.3 Data assimilation and Chesapeake Bay

While the identical twin experiments guide us on the rate of recovery of the control variables, the appropriate density of data, and the definition of an appropriate cost function, data assimilation using tidal elevations (Section 4.2) and real observations (Section 4.3) gives some insight into the physics of the Bay and some empirical quantities used in the model, such as the bottom drag coefficient. Indeed, these results indicate that the bottom drag coefficient displays a periodicity corresponding to the spring-neap tide cycle and its value is a function of the wind speed and direction. The results also confirm Wang's (1979a,b) findings that the response to the wind is different in the lower and upper Bay.

5.3.1 Estimate of the bottom stress and drag coefficient

During the past decades, the bottom stress in tidal models has been defined as a quadratic law with the bottom drag coefficient being determined empirically to

match the amplitude and phase of the observed major tidal constituents. For example, in their study of the coastal seas of southern British Columbia and Washington States, Crean *et al.* (1988) found it necessary to vary spatially the bottom drag coefficient. The overall friction was taken equal to 0.003, with different values (0.03 or 0.006) in the Puget Sound and various straits and passages. In the modeling of tidal motion in the English Channel and southern part of the North Sea, a constant bottom drag coefficient was taken equal to 2.32×10^{-3} (Werner and Lynch, 1987; Ozer and Jamart, 1988). Mofjeld (1988) showed that the quadratic coefficient relating the bottom stress to the vertically averaged velocities depends explicitly on the water depth and found that it can be related to the ratio H/z_o where z_o is the bottom roughness and H the total water depth. Based on the non-rotating channel theory in which frictional drag balances barotropic pressure forcing and shear production of turbulence is balanced by local dissipation (closure of level 2), Mofjeld (1988) found that the bottom drag coefficient is given by

$$c_D = \frac{\kappa^2}{\left[\log(\frac{H}{z_o}) - 1\right]^2} \tag{25}$$

where $\kappa = 0.4$ is the von Karman constant.

Using a variational data assimilation technique and tidal elevations, Lardner *et al.* (1993) estimated the bottom drag coefficient for the Arabian Gulf. The drag coefficient was defined as g/c^2 where $c = C\log(h)$ is the Chezy coefficient, h the undisturbed depth of the water, and g is the acceleration due to gravity. They found in general a higher value for C than the empirical value of 25 used in a previous tidal model for the same region (Lardner *et al.*, 1982). Using an inverse method and tidal-current measurements in the Chesapeake Bay, Bang (1994) found values between 2.0×10^{-4} and 1.6×10^{-3}. Note that all the mentioned studies consider only a spatial variation of the drag coefficient. The temporal variation was not taken into account in 2-D modeling of tidal and wind-driven circulation.

• *Temporal variation of the bottom drag coefficient*

In our study, the bottom drag coefficient was defined in terms of a Chezy coefficient, *i.e.*, it depends on the total depth, and a roughness coefficient called Manning's roughness (Eq. 5). A systematic adjustment of the drag coefficient was done by assimilation of tidal elevations and estimation of the two free parameters, *i.e.*, the exponent of the total depth and the inverse of the Manning's roughness. In the tidal experiment, by assimilating the predicted tidal elevation on 24 hours for 19 consecutive days, temporal variation of the drag coefficient for a spring-neap tide cycle was allowed. Furthermore, the spatial variation was taken into account through the dependence of the coefficient on the total depth. As results of tidal elevations assimilation in the Chesapeake Bay, it is found that the bottom drag coefficient displays a periodicity corresponding to the fortnightly modulation. The drag coefficient varies between 2.5×10^{-4} and 3.1×10^{-3} with a minimum value at spring tide and a maximum at neap tide. While the lower value seems to be smaller than the value found in the literature related to tidal modeling (Johns, 1983; Crean *et al.*, 1988; Ozer and Jamart, 1988), a similar range was found by Bang (1994) in his study in the Chesapeake Bay. The fortnightly modulation is mainly found in the Manning's roughness while the exponent remains close to 1/6, the value commonly used in estuary studies (Officer, 1975).

The first plausible cause for the temporal variation of the drag coefficient can be attributed to the variation of the bottom roughness. Indeed, several studies (McCave, 1973; Taylor and Dyer, 1977; Grant and Madsen, 1982; Davies, 1983; Gross and Nowell, 1983; Wright *et al.*, 1992) showed temporal variability of the bottom drag coefficient due to variation in roughness elements such as ripples and biogenic micromorphology, in movable bed roughness caused by sediment transport and in interactions between waves and currents. Although the temporal changes of the bottom roughness in the Chesapeake Bay are not yet well understood, Wright

et al. (1992) found seasonal variability due to biological processes controlling bed micromorphology and variation at period of a few hours due to wave activity.

A second cause which would account for the periodicity of the bottom drag coefficient can be related to the level of stratification. In their study of the Rhine regions of freshwater influence, Visser *et al.* (1995) showed that the constituents of the tidal currents depend not only on tidal forcing at astronomic periods but also on the less predictable stratification and destratification processes and questioned the robustness of the standard tidal analyses. It has been shown that mixing in the lower Chesapeake Bay and tributaries appears most intense at spring tide while stratification appears most highly developed at neap tide (Haas, 1977; Valle-Levinson, 1995). At spring tide, the vertically averaged current and the bottom current are of the same order of magnitude while the vertically averaged current is smaller than the bottom current at neap tide. Since the bottom stress in our model is defined in terms of the vertically averaged current, the drag coefficient must be larger during neap tide than during spring tide in order to compensate the difference between averaged and bottom current assuming that the bottom drag is the same.

By adjusting the drag coefficient spatially and temporally, the simulation of the tidal circulation in the Chesapeake Bay is improved. Comparable results were found by Lardner *et al.* (1993) in the Arabian Gulf. However, the periodicity of the drag coefficient was not taken into account. As a consequence, the rms error found by Lardner at various stations in the Gulf is higher during spring tide than neap tide (see Fig. 49) which would lead to the proposition that a temporal variation of the drag coefficient is not only true for the Chesapeake Bay, but also for other bodies of water.

Finally, using Eqs. 5 and 25 and an average depth of 8 m, the bottom roughness is found to be equal to 0.01 cm during neap tide which is in good agreement with the value found by Wright *et al.* (1992). During spring tide, the bottom roughness

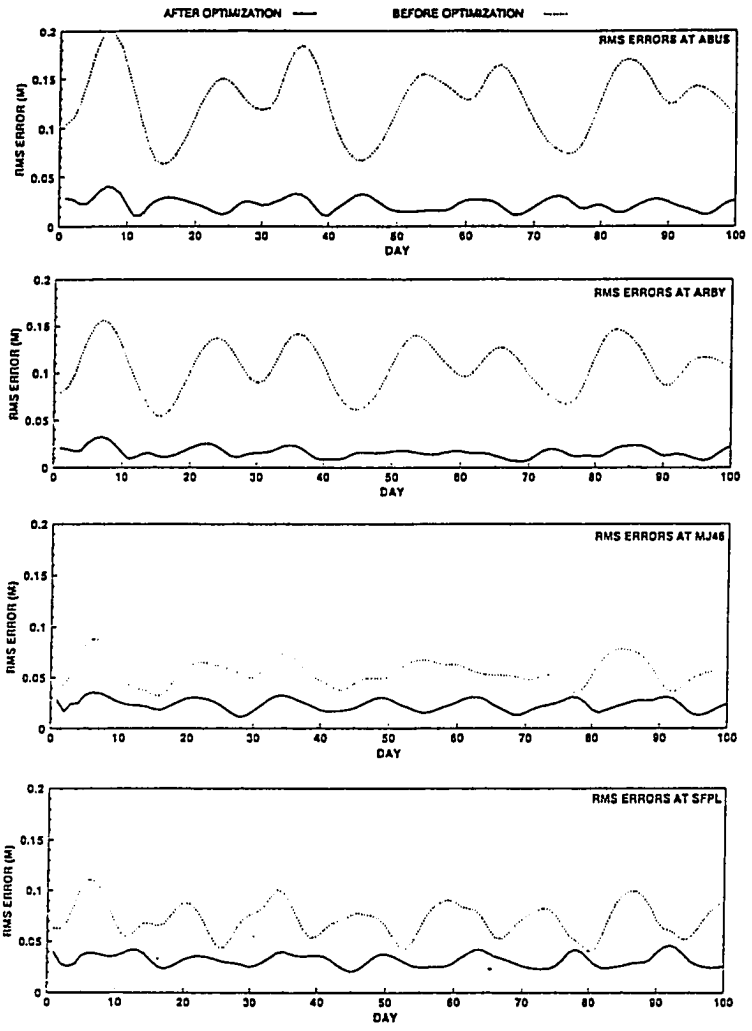

Figure 49: Root-mean square errors at stations in the Arabian Gulf for 100 days following the beginning of the assimilation period. The errors are shown before (dotted line) and after (solid line) optimization. (Lardner *et al.*, 1993).

value averages 0.001 cm which is slightly too small. In order to get a more accurate estimate of the bottom roughness, the relation (25) should be used in the circulation model and z_o be estimated during the assimilation process.

- *Wind effects*

The study of wind-driven circulation shows that the bottom drag coefficient decreased during strong northeasterly and northwesterly wind compared to its value without wind or with a weak wind (Figs. 38, 45). Again, the decrease can be related to the level of stratification and changes in the bottom roughness. During frontal passage, the water column becomes less stratified, indeed destratified (Blumberg and Goodrich, 1990). Therefore, bottom velocity and vertically averaged velocity approach the same value. Furthermore, Wright *et al.* (1992) argued that strong winds should be able to generate 5-s waves large enough to agitate the bed at depth of 10 to 12 m in the Bay and yield variation in the bottom roughness.

Similar wind effect on the bottom friction was found by Ronday (1976) for his study of the North Sea. He showed that a term proportional to the wind stress had to be subtracted from the bottom stress in order to match modeled and observed elevation when the drag coefficient was taken equal to its value without wind. For a northeasterly wind, this corresponds to an increase of the bottom stress which can also be achieved by a decrease of the bottom drag coefficient.

5.3.2 Atmospheric forces in the Bay

The second important feature noticed during the wind-driven circulation experiment is the relative importance of the driving forces in the Bay. While the modeled surface elevation is in excellent agreement with the observed elevations in the main stem, the estimated wind speed and direction in the upper Bay are not in as good agreement with the observations as in the lower Bay. Thomas Point buoy, which is situated

at the narrowest portion of the Bay (Fig. 12) was used to compare the modeled results with the observations in the upper Bay and CBBT tide gauge station was used in the lower Bay. Near Thomas Point, the main stem changes its orientation from northwest-southeast to northeast-southwest and becomes shallower (Fig. 1). Then, during periods of weak wind, the topographic and narrowing effect should be dominant. The model grid spacing, roughly 2 km, is probably too large to correctly resolve the influence of the narrowing of the Bay. Instead, the correction is done to the wind stress in order to minimize the data misfit at Annapolis which is the closest tide gauge station. When the wind becomes stronger, as in the case during a frontal passage, the surface elevation response to the wind is larger and the wind signal in the observations is also stronger. In those conditions, estimated and observed wind speed and direction are in excellent agreement. Further investigation with a finer grid than 2 km is needed to fully investigate the circulation in that area.

A second effect that is neglected and could be a cause for the disagreement between estimated and observed wind is the inverted barometer effect, which induces an increase of the surface elevation for a low pressure system and a decrease of the surface elevation for a high pressure system. For example, Paraso and Valle-Levinson (1995) showed that for 10-11 February 1992, the barometric pressure rise contributed 57% to the sea level change at CBBT and in general the effects of the atmospheric pressure on sea level are not negligible. Vieria (1986) found that the 2-2.5 day sea level oscillations in mid-Bay could not be identified with a seiche in the Bay but could be due to the atmospheric pressure. During the period corresponding to our study, changes in the barometric pressure are also important. From November 1 to November 3, 1990, the atmospheric pressure was about 1023 mb (Fig. 50). It then decreased to a minimum of 1002 mb on November 6 after which it increased to a maximum of 1022 mb in the afternoon of November 7. High pressure during the first three days of November acted to decrease the surface elevation near Thomas

Point and acts against the effect of the wind. The inverted barometer effect, not included in our model, is accounted for through modification of the wind speed and direction.

5.4 Future study

This study represents the first attempt to use surface elevations from tide gauges to estimate the bottom and surface forcings in the Chesapeake Bay. Although the circulation model did not include stratification, river runoff, or the inverted barometer effect and the number of available observations was limited, the modeled surface elevations with the estimated drag coefficient and wind stress are in excellent agreement with the observations at the ten permanent tide gauge stations. Based on the results of the various experiments, several questions however arise. How can the estimated drag coefficient and wind field, and thereby the simulation of the circulation in the Bay, be improved? Can the experiments be repeated for other seasons, *e.g.*, spring and summer? Can the convergence rate of the assimilation process be improved? In order to address these questions, the following investigations are proposed:

- Two remedies should be investigated to improve the estimated wind field in the Bay. Since it has been shown that the inverted barometer effect is not negligible in general (Paraso and Valle-Levinson, 1995), a natural extension of the MU circulation model is to include the inverted barometer effect. The barometric pressure is routinely measured at the major airports, at Thomas Point buoy and at the CBBT station, which could be used to estimate the pressure field over the Bay. Second, assimilation of wind observations from Thomas Point buoy and CBBT station would give more information on the wind field, in addition to surface elevations from the ten tide gauge stations. On the other hand, several twin experiments (not included) showed that as-

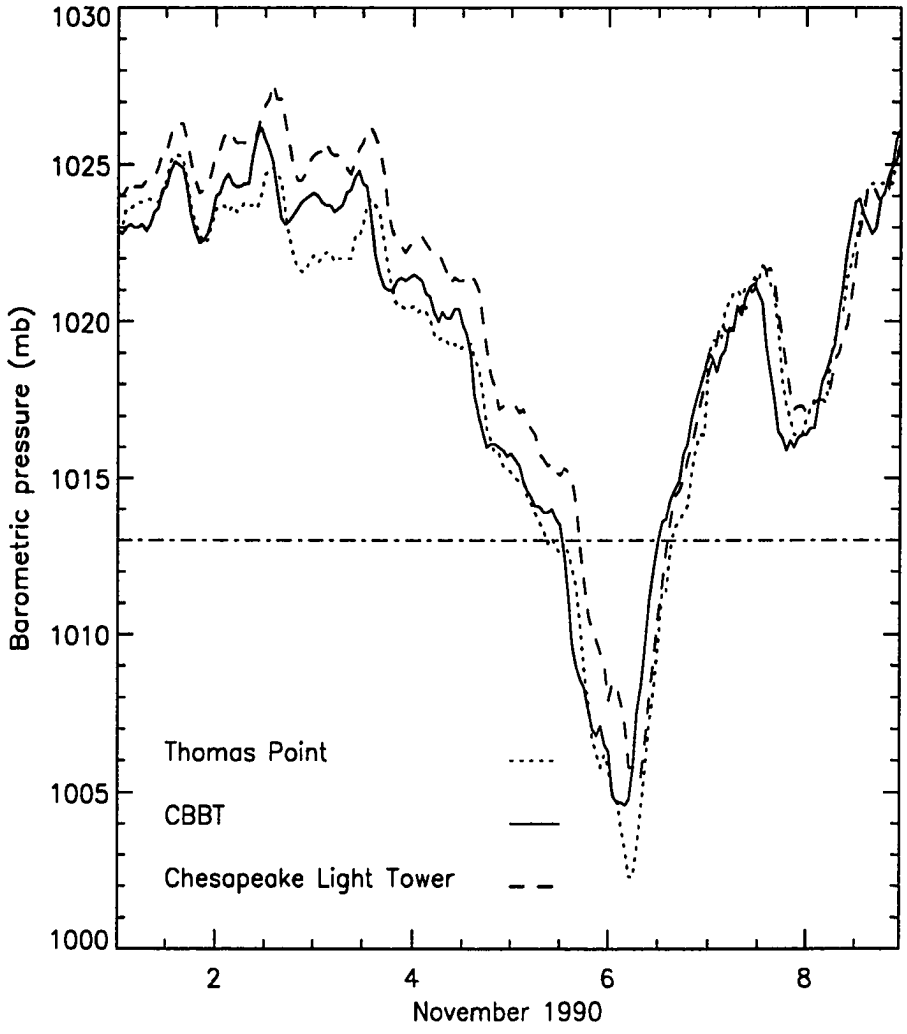

Figure 50: Barometric pressure (mb) from November 1 to November 10, 1990.

similation of surface elevation from additional tide gauges on the eastern shore also gives better estimates of the wind field.

- In our study, the wind field was taken constant during the assimilation period. However, during frontal passages, the wind speed and direction change quite rapidly, as well as the surface elevation in response to it. An estimate of the wind stress every three hours would better represent the temporal changes of the wind field and therefore the wind effect on the circulation in the Bay.

- The bottom stress in the circulation model is taken as a quadratic law with the bottom drag coefficient defined in terms of a Chezy coefficient. Using data assimilation, we might want to estimate the bottom friction in a more general form, *e.g.*, as a free parameter varying in space and time. This would be very important in the case of strong wind when bottom drag coefficient was found to be function of the wind strength (Fig. 45), which could be different from the lower to the upper Bay. Of course, the success of such an experiment will depend on the number of available observations compared to the number of parameters to be estimated (see Section 5.2).

- Data assimilation experiments, shown to be very successful, were conducted for the Fall when the stratification effect is minimal. During spring and summer, stratification may be very important especially at calm wind conditions. Preliminary studies (not included) indicate that, using a two-dimensional model and assimilation of sea level, the surface elevation during times of strong stratification can be estimated with a relative average error of less than 15% at all the stations. A three-dimensional model might however be required at those periods to further improve the estimated circulation in the Bay. Conceptually, data assimilation can be done the same way as for the 2-D modeling. However, a three-dimensional model will require larger memory and computer

time, which could limit the feasibility of the study. Additionally, the number of surface elevation observations might not be sufficient. One could seek other data to assimilate, such as current measurements.

- While the large numbers of iterations needed for convergence in the assimilation process was not an issue in our study, it can become a problem for the aforementioned investigations, when the number of estimated parameters will be larger. Scaling of the parameters, weighting in the cost function (Eq. 8), addition of penalty terms will have to be carefully studied in that case. For example, a preconditioning technique, such as scaling of the gradient of the cost function (Zou and Holloway, 1995) has proved to be very effective in improving the quality of the fit. Furthermore, as pointed out by Tziperman *et al.* (1992a), the issues of scaling and preconditioning should be investigated while considering real observations. The level of noise in the observations might indeed pose a problem in the conditioning of the cost function which would not occur during identical twin experiments.

6 Conclusions

The feasibility of dynamical assimilation of tide gauge observations was investigated to estimate the bottom drag coefficient, the surface stress and the sea level at the Bay mouth, as a first step in improving the modeling of tidal and wind-driven circulation in the Chesapeake Bay. The circulation model used in the study was a 2-D vertically integrated shallow water model where the bottom stress is defined as a quadratic law with a drag coefficient defined in terms of a Chezy coefficient depending on the total depth of the water column and some roughness. The data assimilation technique was the variational adjoint method where the distance between modeled and observed surface elevations is minimized in order to get the optimal value of the control variables. The adjoint model code was developed from the tangent linear code of the circulation model and the optimization technique was the limited memory quasi-Newton method (Gilbert and Lemaréchal, 1989).

Although the model is simple and does not include stratification, river runoff, or inverted barometer effect, the estimate of bottom friction and of surface stress by assimilating tide gauge observations from ten permanent stations yields good agreement between modeled and observed surface elevation in the Bay. It is also found that a one-layer model is adequate to model the sea level and the response to the bottom friction and the wind stress in fall. Whether this is true in spring and summer when the stratification is strong requires further investigations.

The assimilation experiments considered in the present study give some insight into the physics of the Bay as well as into empirical quantities such as the bottom

drag coefficient. It is found that the drag coefficient displays a fortnightly modulation. Its value for depth less than 10 meters doubles from spring to neap tide while the variation is much reduced in deeper regions. This fortnightly modulation is altered by the strength of the wind, which, during a frontal passage, yields a drag coefficient value roughly independent of the depth of the water column. It is also found that the response to meteorological forcing is different in the lower and upper Bay. While the estimated wind field in the lower Bay was in excellent agreement with the wind measured at the Chesapeake Bay Bridge Tunnel, the agreement in the upper Bay was not as good. The disagreement between the estimated wind field in the upper Bay and measured at Thomas Point would indicate that the response of the sea level to the barometric pressure could be as important as the response to the wind forcing in the upper Bay.

From a set of identical twin experiments with model generated data, it was found that a penalty term had to be added to the cost function in order to assure smoothness of the estimated surface elevation field at the Bay mouth. Furthermore, classical scaling of the parameters to bring them to the same order of magnitude was not efficient in accelerating the convergence and yielded a larger error in the estimated parameters.

Finally, in the light of the identical twin, tidal and wind-driven experiments considered in this study, we can conclude that assimilation of tide gauge data in the Chesapeake Bay improved the agreement between modeled and observed surface elevation. We can also propose as a natural extension of this study to include the inverted barometric effect in the model and also repeat the experiments for seasons when stratification is strong.

References

Al-Baali, M., and R. Fletcher, 1986: An efficient line search for nonlinear least squares. *Journal of Optimization Theory and Applications*, **48**, 359–377.

Beale, E. M., 1972: A derivation of conjugate-gradients. *Numerical Methods for Non-linear Optimization*, F.A. Loostma, Ed., Academic Press, 39–43.

Bang, B., 1994: Inverse estimation of horizontal pressure gradients and vertical eddy viscosity profiles in shallow waters. Ph.D. thesis, The School of Marine Science, The College of William and Mary, VA, 154 pp.

Banks, H. T., and K. Kunish, 1989: *Estimation techniques for distributed parameter systems.* Birkhauser, Boston, 315 pp.

Beckers, P. M., and R. J. Neves, 1985: A semi-implicit tidal model of the North European Continental Shelf. *App. Math. Modelling*, **9**, 395–402.

Bennett, A. F., 1985: Array design by inverse method. *Progress in Oceanography*, **15**, Pergamon, 129–156.

Bennett, A. F., and W. P. Budgell, 1987: Ocean data assimilation and the Kalman filter: Spatial regularity. *J. Phys. Oceanogr.*, **17**, 1583–1601.

Bennett, A. F., and P. C. McIntosh, 1982: Open ocean modelling as an inverse problem: Tidal theory. *J. Phys. Oceanogr.*, **12**, 1004–1018.

Bengtsson, L., 1981: *Dynamic meteorology: Data assimilation methods.* M. Ghil and E. Källén, Eds., Springer-Verlag, New York, 330 pp.

Blumberg, A., and D. M. Goodrich, 1990: Modeling of wind-induced destratification in Chesapeake Bay. *Estuaries*,**13**, 236–249.

Bratseth, A. M., 1986: Statistical interpolation by means of successive corrections. *Tellus*, **38A**, 439–447.

Browne, D. R., and C. Fisher, 1988: Tide and tidal currents in the Chesapeake Bay. NOAA Technical Report NOS OMA 3, Rockville, Maryland, 84 pp.

Budgell, W. P., 1986: Nonlinear data assimilation for shallow water equations in branched channel. *J. Geophys. Res.*, **91**, 10633–10644.

Budgell, W. P., 1987: Stochastic filtering of linear shallow water wave processes. *SIAM Journal on Scientific and Statistical Computing*, **8**, 152–170.

Buckley, A., and A. LeNir, 1983: QN-like variable storage conjugate gradients. *Mathematical Programming*, **27**, 155–175.

Cacuci, D. G., 1981: Sensitivity theory for nonlinear system. I. Nonlinear functional analysis approach. *J. Math. Phys.*, **22**, 2794–2802.

Cacuci, D. G., 1988: The forward and adjoint methods of sensitivity analysis. *Uncertainty Analysis*, Yigal Ronen, Ed., CRC Press, Inc., 282 pp.

Cacuci, D. G., and M. C. G. Hall, 1984: Efficient estimation feedback effects with application to climate models. *J. Atmos. Sci.*, **41**, 2063–2068.

Cameron, W. M., and D. W. Pritchard, 1963: Estuaries. *The Sea: Ideas and Observations*, M.N. Hill, Ed., Wiley-Interscience, Vol. **2**, 306–324.

Carrera, J., and S. P. Neumann, 1986a: Estimation of aquifer parameters under transient and steady state conditions. Part I: Maximum likelihood method incorporating prior information. *Water Resources Res.*, **22**, 199–210.

Carrera, J., and S. P. Neumann, 1986b: Estimation of aquifer parameters under transient and steady state conditions. Part II: Uniqueness, stability and solution algorithms. *Water Resources Res.*, **22**, 211–227.

Carrera, J., and S. P. Neumann, 1986c: Estimation of aquifer parameters under transient and steady state conditions. Part III: Application to synthetic and field data. *Water Resources Res.*, **22**, 228–241.

Chuang, W.-S., and W. C. Boicourt, 1989: Seiche motion in the Chesapeake Bay. *J. Geophys. Res.*, **94**, 2105–2110.

Courtier P., 1984: Présentation d'une méthode variationnelle d'assimilation de données metéorologiques réparties dans l'espace et le temps. Note EERM No 101. Available from Meteo-France.

Courtier, P., and O. Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation, II, Numerical results. *Quart. J. Roy. Meteor. Soc.*, **113**, 1329-1347.

Courtier, P., and O. Talagrand, 1990: Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations. *Tellus*, **42A**, 531–549.

Courtier P., J. Derber, R. Errico, J-F Louis, and T. Vukicevic, 1993: Important literature on the use of adjoint, variational methods and the Kalman filter in meteorology. *Tellus*, **45A**, 342–357.

Crean, P. B., T. S. Murty, and J. A. Stronach, 1988: Mathematical modeling of tides and estuarine circulation, the Coastal Seas of Southern British Columbia and Washington State. *Lecture Notes on Coastal and Estuarine Studies*, M. J Bowman *et al.*, Eds., Springer-Verlag, Vol. **30**, 471 pp.

Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367–374.

Das, S. K., and R. W. Lardner, 1991: On the estimation of parameters of hydraulic models by assimilation of periodic tidal data. *J. Geophys. Res.*, **96**, 15187–15196.

Das, S. K., and R. W. Lardner, 1992: Variational parameter estimation for a two-dimensional numerical tidal model. *Int. J. Numer. Methods Fluids*, **15**, 313–327.

Davidon, W. C., 1959: Variable metric methods for minimization. A.E.C. Research and development report ANL-5990, Argonne National Laboratory, Argonne, Il.

Davies, A. G., 1983: Wave interactions with rippled sand beds. *Physical Oceanography of Coastal and Shelf Seas*, B. Johns, Ed., Elsevier, 1–64.

Defant, A., 1961. *Physical Oceanography*, Vol. 2., Pergamon Press, New York, NY, 598 pp.

Dembo, R. S., and T. Steibaug, 1983: Truncated-Newton algorithms for large-scale unconstraint optimization. *Mathematical Programming*, **26**, 190–212.

Derber, J. C., 1985: The variational four-dimensional assimilation analyses using filtered models as constraints. Ph.D. dissertation, University of Wisconsin, Madison, WI.

Dyer, K. R., 1973: *Estuaries: a Physical Introduction.* John Wiley and Sons, London, England, 140 pp.

Farrell, B. F., and A. M. Moore, 1992: An adjoint method for obtaining the most rapidly growing perturbation to oceanic flows. *J. Phys. Oceanogr.*, **22**, 338–349.

Fisher, C. W., 1986: Tidal circulation in Chesapeake Bay. Ph.D. dissertation, Old Dominion University, Norfolk, VA, 255 pp.

Fletcher, R., 1987: *Practical Methods of Optimization.* Second Edition, John Wiley and Sons, New York, 436 pp.

Fukumori, I., J. Benveniste, C. Wunsch, and D. B. Haidvogel, 1993: Assimilation of sea surface topography into an ocean circulation model using a steady-state smoother. *J. Phys. Oceanogr.*, **23**, 1831–1855.

Gandin, L. S., 1963: *Objective analysis of meteorological fields.* (From the Russian) Israel Program for Scientific Translations, 1965, Jerusalem, 242 pp.

Ghil, M., and P. Malanotte-Rizzoli, 1991: Data assimilation in meteorology and oceanography. *Advances in Geophysics*, **33**, 141–266.

Ghil, M., S. Cohn, J. Taqvantzis, K. Bube, and E. Isaacson, 1981: Applications of estimation theory to numerical weather prediction. *Dynamic Meteorology: Data Assimilation Methods*, L. Bengtsen, M. Ghil and E. Kallen, Eds., Springer-Verlag, New York, 139–224, 330 pp.

Giering, R., 1995: Adjoint code generation. Max-Planck-Institute für Meteorologie. In preparation.

Gilbert J.-Ch., and C. Lemaréchal, 1989: Some numerical experiments with variable-storage quasi-Newton algorithms. *Mathematical Programming*, **45**, 407–435.

Gill, P., E. W. Murray, and M. H. Wright, 1981: *Practical Optimization.* Academic Press, Inc., London, 401 pp.

Goodrich, D. M., 1985: On stratification and wind-induced mixing in Chesapeake Bay. Ph.D. dissertation, State University of New York, Stony Brook, New York, 134 pp.

Goodrich, D. M., W. C. Boicourt, P. Hamilton, and D. W. Pritchard, 1987: Wind-induced destratification in Chesapeake Bay. *J. Phys. Oceanogr.*, **17**, 2232–2240.

Grant, W.D., and O. S. Madsen, 1982: Movable bed roughness in oscillatory flow. *J. Geophys. Res.*, **87**, 469–481.

Greiner, E., and C. Périgaud, 1994: Assimilation of Geosat altimetric data in a nonlinear reduced-gravity model of the Indian Ocean. Part I: Adjoint approach and model-data consistency. *J. Phys. Oceanogr.*, **24**, 1783–1804.

Gross, T. F., and A. R. M. Nowell, 1983: Mean flow and turbulence scaling in a tidal boundary flow. *Contin. Shelf Res.*, **2**, 109–126.

Haas, L. W., 1977: The effect of the spring-neap tidal cycle on the vertical salinity structure of the James, York and Rappahannock Rivers, Virginia, U.S.A. *Estuarine and Coastal Marine Science*, **5**, 485–496.

Haidvogel, D. B., and A. R. Robinson, Eds, 1989: Special issue: data assimilation. *Dynamics of the Atmosphere and Oceans*, **13**, 171–513.

Haight, F. J., H. E. Finnegan, and G. L. Anderson, 1930: Tides and currents in Chesapeake Bay and tributaries. Special publication No. 162. U.S. Coast and Geodetic Survey, Washington, D.C., 143 pp.

Hall, M. C. G., and D. G. Cacuci, 1983: Physical interpretation of the adjoint functions for sensitivity analysis of atmospheric models. *J. Atmos. Sci.*, **40**, 2537–2546.

Hall, M. C. G., D. G. Cacuci, and M. E. Schlesinger, 1982: Sensitivity analysis of a radiative-convective model by adjoint method. *J. Atmos. Sci.*, **39**, 2038–2050.

Harris, R. A., 1907: Currents, shallow water tides, meteorological tides and miscellaneous matters. *Manual of Tides*, Appendix 6, Chapter 6:357, Annual Report for 1907. U.S. Coast and Geodetic Survey, Washington, D.C.

Heeminck, A. W., and H. Kloosterhuis, 1990: Data assimilation for non-linear tidal models. *Int. J. Numer. Methods Fluids*, **11**, 1097–1112.

Hicks, S. D., 1964: Tidal wave characteristics of Chesapeake Bay. *Chesapeake Science*, **5** (3), 103–113.

Hoffman, R. N., 1986: A four-dimensional analysis exactly satisfying equations of motion. *Mon. Wea. Rev.*, **114**, 388–397.

Jamart, B. M., and J. Ozer, 1989: Some results and comments on the Tidal Flow exercise. *Advances in Water Resources*, **12**, 211–220.

Johns, B., 1983: Turbulence modelling beneath waves over beaches. *Physical Oceanography of Coastal and Shelf Seas*, B. Johns, Ed., Elsevier, 111–133.

Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. *Journal of Basic Engineering (Transactions of the ASME)*, **82D**, 35–45.

Kalman, R. E., and R. S. Bucy, 1961: New results in linear filtering and prediction theory. *Journal of Basic Engineering (Transactions of the ASME)*, **83D**, 95–108.

Kravaris, C., and J. H. Seinfeld, 1985: Identification of parameters in distributed parameters systems by regularization. *SIAM J. on Control and Optimization*, **23**, 217–241.

Kunish, K., and L. White, 1986: Parameter estimation for elliptic equations in multidimensional domain with point and flux observations. *Nonlinear Analysis, Theory, Methods and Applications*, **10**, 121–146.

Lacarra, J., and O. Talagrand, 1988: Short-range evolution of small perturbations in a barotropic model. *Tellus*, **40A**, 81–95.

Lardner, R. W., 1993: Optimal control of open boundary conditions for a numerical tidal model. *Comput. Methods Appl. Mech. Eng.*, **102**, 367–387.

Lardner, R. W., and S. K. Das, 1994: Optimal estimation of eddy viscosity for a quasi-three-dimensional numerical tidal and storm surge model. *Int. J. Numer. Methods Fluids*, **18**, 295–312.

Lardner, R. W., A. H. Al-Rabeh, and N. Gunay, 1993: Optimal estimation of parameters for a two-dimensional hydrodynamical model of the Arabian Gulf. *J. Geophys. Res.*, **98**, 18,229–18,242.

Lardner, R. W., M. S. Belen, and H. M. Cekirge, 1982: Finite difference model for tidal flows in the Arabian Gulf. *Comp. Math. Appl.*, **8**, 425–444.

Lawson, L. M., E. E. Hofmann, and Y. H. Spitz, 1995: Time series sampling and data assimilation in a simple marine ecosystem model. Submitted.

Lawson, L. M., Y. H. Spitz, E. E. Hofmann, and R. B. Long, 1995: A data assimilation technique applied to a predator-prey model. *Bulletin of Mathematical Biology*, **57**, 593–617.

Le Dimet, F. X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus*, **38A**, 97–110.

Legler, D. M., and I. M. Navon, 1991: VARIATM – A Fortran code for objective analysis of pseudo-stress with large-scale conjugate-gradient minimization. *Computers and Geosciences*, **17(1)**, 1–21.

Legler, D. M., I. M. Navon, and J.J. O'Brien, 1989: Objective analysis of pseudo-stress over the Indian Ocean using a direct minimization approach. *Mon. Wea. Rev.*, **117**, 709–720.

Lemaréchal, C., 1981: A view of line-searches. *Optimization and Optimal Control*, Lecture Notes in Control and Information Science, A. Auslender, W. Oettli, J. Stoer, Eds., Springer-Verlag, Heidelberg, 59–78.

Lewis, J. M., and J. C. Derber, 1985: The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, **37A**, 309–322.

Li Y., I. M. Navon, P. Courtier, and P. Gauthier, 1993: Variational data assimilation with a semi-implicit semi-Lagrangian global shallow-water equations model and its adjoint. *Mon. Wea. Rev.*, **121**, 1759–1769.

Liu, D. C., and J. Nocedal, 1989: On the limited memory BFGS method for large scale minimization. *Math. Prog.*, **45**, 503–528.

Long, R. B., and W. C. Thacker, 1989a: Data assimilation into a numerical equatorial ocean model, Part I: the model and the assimilation algorithm. *Dyn. Atmos. Oceans*, **13(3-4)**, 379–412.

Long, R. B., and W. C. Thacker, 1989b: Data assimilation into a numerical equatorial ocean model, Part II: assimilation experiments. *Dyn. Atmos. Oceans*, **13(3-4)**, 413–440.

Lorenc, A. C., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109**, 701–721.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112**, 1177–1194.

McCave, I. N., 1973: Some boundary-layer characteristics of tidal currents bearing sand in suspension. *Mémoires Société Royale des Sciences de Liège*, **6**, 187–206.

Miller, R. N., 1986: Towards the application of the Kalman filter to regional open ocean modeling. *J. Phys. Oceanogr.*, **16**, 72–86.

Mofjeld, H. O., 1988: Depth dependence of bottom stress and quadratic coefficient for barotropic pressure-driven currents. *J. Phys. Oceanogr.*, **18**, 1658–1669.

Moore, A. M., 1991: Data assimilation in a quasi-geostrophic open-ocean model of the Gulf Stream region using the adjoint method. *J. Phys. Oceanogr.*, **21**, 398–427.

Morse, P., and H. Feshbach, 1953: *Method of Theoretical Physics. Part I.* McGraw-Hill, New York.

Nash, S. G., 1984a: Newton-type minimization via the Lanczos method. *SIAM J. Numer. Anal.*, **21**, 770–788.

Nash, S. G., 1984b: Solving nonlinear programming problems using truncated-Newton techniques. *Numerical optimization*, P.T. Boggs, R.H. Byrd and R.B. Schnabel, Eds., SIAM, Philadelphia, 119–136.

Navon, I. M., 1986: A review of variational and optimization methods in meteorology. *Variational Methods in Geosciences*, Y. Sasaki, Ed., Elsevier, New York, 29–34.

Navon, I.M., and R. de Villiers, 1983: Combined penalty multiplier optimization methods to enforce integral invariants conservation. *Mon. Wea. Rev.*, **111**, 1228–1243.

Navon, I. M., and D. M. Legler, 1987: Conjugate-gradient methods for large-scale minimization in meteorology. *Mon. Wea. Rev.*, **115**, 1479–1502.

Navon, I. M., X. Zou, J. Derber, and J. Sela, 1992a: Variational data assimilation with an adiabatic version of the NMC spectral model. *Mon. Wea. Rev.*, **120**, 1433–1446.

Navon, I. M., X. Zou, M. Berger, P. K. H. Phua, T. Schlick, and F. X. Le Dimet, 1992b: Numerical experience with limited memory quasi-Newton and truncated Newton methods. *Optimization Techniques and Applications*, K.H Phua *et al.*, Eds., World Scientific Publishing Co., Vol. **1**, 33–48.

Navon, I. M., X. Zou, M. Berger, P. K. H. Phua, T. Schlick, and F. X. Le Dimet, 1992c: Testing for reliability and robustness of optimization codes for large scale optimization problems. *Optimization Techniques and Applications*, K.H Phua *et al.*, Eds., World Scientific Publishing Co., Vol. **1**, 445–480.

Nocedal, J., 1980: Updating quasi-Newton matrices with limited storage. *Mathematics of Computation*, **35**, 773–782.

Officer, C. B., 1976: *Physical Oceanography of Estuaries (and Associated Coastal Waters)*. John Wiley and Sons, 465 pp.

O'Leary, D. P., 1982: A discrete Newton algorithm for minimizing a function of many variables. *Math. Progr.*, **23**, 20–33.

Omatu, S., and J. H. Seinfeld, 1989: *Distributed parameter systems: Theory and applications*. Oxford Mathematical Monographs. Oxford University Press, 430 pp.

Ozer, J., and B. M. Jamart, 1988: Tidal motion in the English Channel and southern North Sea : Comparison of various observational and model results. *Computational Methods in Water Resources, Vol. 1 : Modeling Surface and Sub-surface Flows*, M.A. Celia, L.A. Ferrand, C.A. Brebbia, W.G. Gray and G.F. Pinder, Eds., Computational Mechanics Publications, Elsevier, 267–273.

Ozer, J., E. Deleersnijder, and B. M. Jamart, 1990: Model intercomparison : description of a semi-implicit numerical model for shallow-water wave equations. MUMM's contribution to MAST-0050-C (SMA), Technical Report 1.

Panchang, V. G., and J. J. O'Brien, 1989: On the determination of hydraulic model parameters using the adjoint state formulation. *Modelling in Marine Systems*, A.M. Davies, Ed., CRC Press, Boca Raton, 6–18.

Paraso, M., and A. Valle-Levinson, 1995: Meteorological influences on sea level and water temperature in the lower Chesapeake Bay: 1992. Submitted.

Pickard, G., and W. J. Emery, 1982: *Descriptive Physical Oceanography: An Introduction.* Pergamon Press, 4th edition, 249pp.

Polak, E., and G. Ribiere, 1969: Note sur la convergence des méthodes de directions conjugées. *Rev. Franc. Informat. Rech. Operationnelle*, **16**, 35–43.

Powell, M. J. D., 1977: Restart procedures of the conjugate gradient method. *Mathematical Programming*, **12**, 241–254.

Powell, M. J. D., 1985: Convergence properties of algorithms for nonlinear optimization. Numerical Analysis Report 1985/NA1, DAMTP, Silver Street, Cambridge CB3 9EW, England.

Pritchard, D. W., 1952a: Estuarine hydrography. *Advan. Geophy.*, **1**, 243–280.

Pritchard, D. W., 1952b: Salinity distribution and circulation in the Chesapeake Bay estuaries system. *J. Mar. Res.*, **11**, 106–123.

Pritchard, D. W., and S. R. Rives, 1979: Physical hydrography and dispersion in a segment of the Chesapeake Bay adjacent to the Calvert Cliffs nuclear power plant. Chesapeake Bay Institute, Special Report, No. 74.

Rennie, S. E., and B. Neilson, 1994: Chesapeake Bay Atlas US EPA Monitoring Program Water Quality Data 1984-1991, Virginia Institute of Marine Science, Data Report No. 55, 13 pp.

Richardson, J. E., and V. G. Panchang, 1992: A modified adjoint method for inverse eddy viscosity estimation in coastal circulation models. *Estuarine and Coastal Modeling, Proc. 2nd Int. Conf.*, M. L. Spaulding, Ed., Amer. Soc. Civil Engrs., New York, 733–745.

Ronday, Fr., 1976: Modèles Hydrodynamiques. *Modélisation des systèmes marins. Projet Mer. Rapport Final*, J.C.J. Nihoul, Ed., Services du Premier Ministre, Programmation de la Politique Scientifique, Bruxelles, Vol. 3, 270 pp.

Sasaki, Y., 1955: A variational study of the numerical prediction based on the variational principle. *J. Meteor. Soc. Japan*, **33**, 262–275.

Sasaki, Y., 1970: Some basic formalisms in numerical variational analysis. *Mon. Wea. Rev.*, **98**, 875–883.

Schröter, J., and C. Wunsch, 1986: Solution of nonlinear finite difference ocean models by optimization methods with sensitivity and observational strategy analysis. *J. Phys. Oceanogr.*, **16**, 1855–1871.

Schlick, T., and A. Fogelson, 1992a: TNPACK- A Truncated Newton minimization package for large-scale problems: I. Algorithm and usage. ACMTOMS, **18** (1), 46–70.

Schlick, T., and A. Fogelson, 1992b: TNPACK- A Truncated Newton minimization package for large-scale problems: II. Implementation examples. ACMTOMS, **18** (1), 71–111.

Seiler, U., 1993: Estimation of open boundary conditions with the adjoint method. *J. Geophys. Res.*, **98**, 22,855–22,870.

Shanno, D. F., 1978: Conjugate gradient methods with inexact searches. *Meth. Oper. Res.*, **3**, 244–256.

Shanno, D. F., and K. H. Phua, 1976: Algorithm 500, minimization of unconstraint multivariate functions. *ACM Transactions on Mathematical Software*, **2**, 87–94.

Sheinbaum, J., and D. L. T. Anderson, 1990: Variational assimilation of XBT data. Part I. *J. Phys. Oceanogr.*, **20**, 672–688.

Smedstad, O. M., and J. J. O'Brien, 1991: Variational data assimilation and parameter estimation in an equatorial Pacific ocean model. *Prog. Oceanog.*, **26**, 179-241.

Schwab, D. J., 1982: An inverse method for estimating wind stress from water-level fluctuations. *Dyn. Atmos. Oceans*, **6**, 251-278.

Talagrand, O., 1991: The use of adjoint equations in numerical modeling of the atmospheric circulation. *SIAM , Automatic Differentiation: Theory, Implementation, and Application*, A. Griewanek and G. Corliss, Eds., SIAM, Philadelphia, PA, 169–180.

Talagrand, O., and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. Part I: Theory. *Quart. J. Roy. Meteor. Soc.*, **113**, 1311–1328.

Taylor, P. A., and K. R. Dyer, 1977: Theoretical models of flow near the bed and their implication for sediment transport. *The Sea*, Vol. **6**, E.D. Goldberg, I.N. McCave, J.J. O'Brien, and J.H. Steele, Eds., Interscience, New York, 579–601.

Thacker, W. C., 1987: Three lectures on fitting numerical models to observations. *External Rep., GKSS 87/E/65*, GKSS Forschungszentrum Geesthacht, Geesthacht, Federal Republic of Germany.

Thacker, W. C., and R. B. Long, 1988: Fitting dynamics to data. *J. Geophys. Res.*, **93**, 1227–1240.

Thépaut, J. N., and P. Courtier, 1991: Four-dimensional data assimilation using the adjoint of a multilevel primitive equation model. *Quart. J. Roy. Meteor. Soc.*, **117**, 1225–1254.

Thépaut, J. N., D. Vasiljevic, P. Courtier, and J. Pailleux, 1993: Variational assimilation of conventional meteorological observations with a multilevel primitive equation model. *Quart. J. Roy. Meteor. Soc.*, **119**, 153–186.

Toint, Ph. L., 1981: Towards an efficient sparsity exploiting Newton method for minimization. *Sparse Matrices and their Uses, The Institute of Mathematics and its Applications Conference Series*, I.S. Duff, Ed.,Academic Press, New York, 57–87, 387 pp.

Tziperman, E., W. C. Thacker, R. B. Long, and S. -M. Hwang, 1992a: Oceanic data analysis using a general circulation model. Part I: Simulations. *J. Phys. Oceanogr.*, **22**, 1434–1457.

Tziperman, E., W. C. Thacker, R. B. Long, S. -M. Hwang, and S. R. Rintoul, 1992b: Oceanic data analysis using a general circulation model. Part II: A North Atlantic model. *J. Phys. Oceanogr.*, **22**, 1458–1485.

Valle-Levinson, A., 1995: Observation of barotropic and baroclinic exchanges in the lower Chesapeake Bay. *Contin. Shelf Res.*, in press.

Vieira, M., 1986: The meteorologically driven circulation in mid-Chesapeake Bay. *J. Mar. Res.*, **44**, 473–493.

Visser, A. W., A. J. Souza, K. Hessner, and J. H. Simpson, 1995: The effect of stratification on tidal current profiles in a region of freshwater influence. Oceanologica Acta, in press.

Wang, D. -P., 1979a: Subtidal sea level variations in the Chesapeake Bay and relations to atmospheric forcing. *J. Phys. Oceanogr.*, **9**, 413–421.

Wang, D. -P., 1979b: Wind-driven circulation in the Chesapeake Bay, winter 1975. *J. Phys. Oceanogr.*, **9**, 564–572.

Wang, D. -P., and A. Elliott, 1978: Non-tidal variability in the Chesapeake Bay and Potomac River: Evidence for non-local forcing. *J. Phys. Oceanogr.*, **8**, 225–232.

Wang, Z., 1993: Variational data assimilation with 2-D shallow water equations and 3-D FSU global spectral models. Tech. Rep. FSU-SCRI-93T-149, Florida State University, Tallahassee, Florida, 235 pp.

Wang, Z., I. M. Navon, F. X. Le Dimet, and X. Zou, 1992: The second order adjoint analysis: Theory and application. *Meteorology and Atmospheric Physics* , **50**, 3–20.

Wang, Z., I. M. Navon, X. Zou, and F.X. Le Dimet, 1995: A truncated-Newton optimization algorithm in meteorology applications with analytic Hessian/vector products. *Computational Optimization and Applications*, Vol. **4**, 241–262.

Weisberg, R. H., 1976: The nontidal flow in the Providence River of Narragansett Bay: A stochastic approach for estuarine circulation. *J. Phys. Oceanogr.*, **6**, 345–354.

Werner, F. E., and D. R. Lynch, 1987: Field verification of Wave Equation tidal dynamics in the English Channel and southern North Sea. *Advances in Water Resources*, **10**, 115–130.

Wolf, P., 1969: Convergence conditions for ascent methods. *SIAM Review*, **11**, 226–235.

Wong, K. -C., and R. Garvine, 1984: Observations of subtidal, wind-induced variability in the Delaware estuary. *J. Geophys. Res.*, **89**, 10,589–10,597.

Wright, L. D., J. D. Boon, J. P. Xu, and S. C. Kim, 1992: The bottom boundary layer of the bay stem plains environment of lower Chesapeake Bay. *Estuarine Coastal Shelf Sc.*, **35**, 17–36.

Yeh, W. W.-G., 1986: Review of parameter estimation procedures in groundwater hydrology. *Water Resources Res.*, **22**, 95–108.

Yu, L., and J. J. O'Brien, 1991: Variational estimation of the wind stress drag coefficient and the oceanic eddy viscosity profile. *J. Phys. Oceanogr.*, **5**, 709–719.

Zou, J., and G. Holloway, 1994: Improving steady-state fit of dynamics to data using adjoint equation for gradient preconditioning. *Mon. Wea. Rev.*, **13**, 199–211.

Zou, X., I. M. Navon, and F. X. Le Dimet, 1992a: Incomplete observations and control of gravity waves in variational data assimilation. *Tellus*, **44A**, 273–296.

Zou, X., I. M. Navon, and F. X. Le Dimet, 1992b: An optimal nudging data assimilation scheme using parameter estimation. *Quart. J. Roy. Meteor. Soc.*, **118**, 1163–1186.

Zou, X., A. Barcilon, I. M. Navon, J. Whitaker, and D. G. Cacuci, 1993b: An adjoint sensitivity study of blocking in a two-layer isentropic model. *Mon. Wea. Rev.*, **12**, 2833–2857.

Zou, X., I. M. Navon, M. Berger, K. H. Phua, T. Schlick, and F. X. Le Dimet, 1993a: Numerical experience with limited-memory quasi-Newton and truncated Newton methods. *SIAM Journal on Optimization*, **3**, 582–608.

Appendices

A Construction of the adjoint code

A.1 Tangent linear model method

The following simple example shows how to construct the adjoint code from the tangent linear code. We also show that the gradient of the cost function is obtained from the adjoint code without extra computation. Consider the following set of equations that can be thought of as the discretized model equations.

$$\begin{aligned} y_1 &= 2x_1 \\ y_2 &= x_2 - x_1 \\ y_3 &= y_1 y_2 \\ J &= \tfrac{1}{2}(y_3 - \hat{y})^2 \end{aligned} \tag{A1}$$

One can think of this model in steps

$$(x_1, x_2) \Longrightarrow [G_1] \Longrightarrow (y_1, y_2) \Longrightarrow [G_2] \Longrightarrow (y_3)$$

In terms of the notations used in Eq. (9), $\mathbf{Y} = G_2 G_1(\mathbf{X}), \delta\mathbf{Y} = G_2' G_1'(\delta\mathbf{X})$, where $G' = G_2' G_1'$ so that $(G')^* = (G_1')^* (G_2')^*$. For the example

$$G_1' = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix},$$

$$G_2' = \begin{bmatrix} \frac{\partial y_3}{\partial y_1} & \frac{\partial y_3}{\partial y_2} \end{bmatrix} = \begin{bmatrix} y_2 & y_1 \end{bmatrix}.$$

Hence using $\mathbf{Y}_1 = (y_1, y_2)$ and $\mathbf{Y}_2 = (y_3)$ we have

$$\nabla_{\mathbf{X}} J = (G_1')^* (G_2')^* \nabla_{\mathbf{Y}_2} J$$

or

$$\nabla_{\mathbf{X}} J = \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_2 \\ y_1 \end{bmatrix} [y_3 - \hat{y}] = \begin{bmatrix} 2y_2 - y_1 \\ y_1 \end{bmatrix} [y_3 - \hat{y}]$$

which gives

$$\nabla_{\mathbf{X}} J = \begin{bmatrix} (2y_2 - y_1)(y_3 - \hat{y}) \\ y_1(y_3 - \hat{y}) \end{bmatrix} \tag{A2}$$

The tangent linear model from (A1) is

$$\begin{aligned} \delta y_1 &= 2\delta x_1 \\ \delta y_2 &= -\delta x_1 + \delta x_2 \\ \delta y_3 &= y_1 \delta y_2 + y_2 \delta y_1 \end{aligned} \tag{A3}$$

Consider each line of the tangent linear model separately and in reverse order. The last line can be written in a matrix form as

$$\begin{bmatrix} \delta y_1 \\ \delta y_2 \\ \delta y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ y_2 & y_1 \end{bmatrix} \begin{bmatrix} \delta y_1 \\ \delta y_2 \end{bmatrix}$$

Its adjoint can be written out by matrix transposition as

$$\begin{bmatrix} ay_1 \\ ay_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & y_2 \\ 0 & 1 & y_1 \end{bmatrix} \begin{bmatrix} ay_1 \\ ay_2 \\ ay_3 \end{bmatrix}$$

or

$$\begin{aligned} ay_1 &= ay_1 + y_2 ay_3 \\ ay_2 &= ay_2 + y_1 ay_3 \end{aligned} \tag{A4}$$

if we replace δ by a to denote an adjoint variable. By repeating the same sequence of operations for the second and first line of the tangent linear model respectively, we obtain

$$\begin{aligned} ax_1 &= ax_1 - ay_2 \\ ax_2 &= ax_2 + ay_2 \end{aligned} \tag{A5}$$

and

$$ax_1 = ax_1 + 2ay_1 \tag{A6}$$

Finally, from the definition of the cost function (last equation of (A1)), we can obtain the forcing to apply to the adjoint system, *i.e.*,

$$ay_3 = (y_3 - \hat{y}) \tag{A7}$$

Combining (A4), (A5), (A6), (A7), the adjoint model corresponding to the non-linear model (A1) is

$$\begin{aligned}
&ay_1 = ay_2 = ay_3 = ax_1 = ax_2 = 0 \\
&ay_3 = (y_3 - \hat{y}) \\
&ay_2 = ay_2 + y_1 ay_3 \\
&ay_1 = ay_1 + y_2 ay_3 \\
&ax_1 = ax_1 - ay_2 \\
&ax_2 = ax_2 + ay_2 \\
&ax_1 = ax_1 + 2ay_1
\end{aligned} \tag{A8}$$

where y_1, y_2, y_3 are given by the non-linear model. Note that combining (A8) in order gives,

$$\begin{aligned}
ax_2 &= y_1(y_3 - \hat{y}) \\
ax_1 &= 2y_2(y_3 - \hat{y}) - y_1(y_3 - \hat{y}) \quad = (2y_2 - y_1)(y_3 - \hat{y})
\end{aligned} \tag{A9}$$

A comparison of (A2) and (A9) yields

$$ax_1 = \frac{\partial J}{\partial x_1}$$
$$ax_2 = \frac{\partial J}{\partial x_2}$$

The "tangent linear technique" gives then a systematic way to obtain the gradient of the cost function with respect to control variables.

It is worth noting that the adjoint equations have to be written in reverse order from the tangent linear model. The adjoint variables have to be carefully initialized

to zero at the beginning of the adjoint code. The adjoint variables are accumulated for every equation where the direct variables appear on the right side of the equation. Finally, the quantities computed during the direct computation are used in the adjoint computation. Therefore, they have to be stored during the direct computation.

With that simple example, we have shown that the writing of the adjoint code from the code is a straightforward operation. To any subroutine of the direct code will correspond an adjoint subroutine. Furthermore, to any statement in one subroutine of the direct code will correspond an statement in the adjoint code.

A.2 Lagrange multiplier technique

In this section, we show how to apply the Lagrange multiplier technique developed in Section 3.4.2 and how to write the adjoint equations for the model (A1). For the example (A1), the Lagrange function defined in Eq. 14 is

$$L = J - \lambda_J(J - \frac{1}{2}(y_3 - \hat{y})^2) - \lambda_{y_1}(y_1 - 2x_1) - \lambda_{y_2}(y_2 - x_2 + x_1) - \lambda_{y_3}(y_3 - y_1 y_2).$$

Requiring that the derivatives of L vanish with respect to J, y_3, y_2, y_1 yields

$$1 - \lambda_J = 0 \Longrightarrow \lambda_J = 1$$

$$\lambda_J(y_3 - \hat{y}) - \lambda_{y_3} = 0 \Longrightarrow \lambda_{y_3} = (y_3 - \hat{y})\lambda_J$$

$$-\lambda_{y_2} + \lambda_{y_3} y_1 = 0 \Longrightarrow \lambda_{y_2} = y_1 \lambda_{y_3}$$

$$-\lambda_{y_1} + \lambda_{y_3} y_2 = 0 \Longrightarrow \lambda_{y_1} = y_2 \lambda_{y_3}$$

Requiring that the derivatives of L vanish with respect to x_2, x_1 and letting the gradient be stored in the Lagrange multipliers associated with x_2, x_1 yields

$$\lambda_{x_2} = \lambda_{y_2}$$

$$\lambda_{x_1} = 2\lambda_{y_1} - \lambda_{y_2}$$

Note that the last equation contains contributions from the first and second equation of the model (A1).

If we replace λ by a to denote an adjoint variable corresponding to the model and control variables, the adjoint code corresponding to the system of equations (A1) is

$$\begin{aligned} & ay_3 = ay_2 = ay_1 = ax_2 = ax_1 = 0 \\ & aJ = 1 \\ & ay_3 = (y_3 - \hat{y})aJ \\ & ay_2 = y_1 ay_3 + ay_2 \\ & ay_1 = y_2 ay_3 + ay_1 \\ & ax_2 = ay_2 + ax_2 \\ & ay_1 = -ay_2 + ax_1 \\ & ax_1 = 2ay_1 + ax_1 \end{aligned} \tag{A10}$$

It should be pointed out that there is no need to write the Lagrange function. The adjoint model code is simply written by considering the direct code line-by-line in reverse order. To every variable of the right-hand side of any line of direct code corresponds an adjoint variable and therefore an adjoint equation. The adjoint variables have to be carefully initialized to zero at the beginning of the adjoint code.

Notice that combining (A10) in order gives

$$\begin{aligned} & ax_2 = -y_1(y_3 - \hat{y}) \\ & ax_1 = 2y_2(y_3 - \hat{y}) - y_1(y_3 - \hat{y}) = (2y_2 - y_1)(y_3 - \hat{y}) \end{aligned}$$

and from (A2) we have

$$\begin{aligned} ax_1 &= \frac{\partial J}{\partial x_1} \\ ax_2 &= \frac{\partial J}{\partial x_2} \end{aligned}$$

This method gives a straightforward procedure to derive the adjoint code and

compute the gradient of the cost function without having to derive the tangent linear code which can be a tedious task for a complicated model. There is therefore less chance of error coding. However, it can have its disadvantage in the fact that the gradient of the cost function can only be checked at the end of the coding. If an error is introduced in the coding, it can be hard to know in which subroutine the mistake was introduced. By contrast, as shown in Section 3.4.3, each subroutine of the adjoint code can be checked separately when using the tangent linear method.

B Optimization algorithms

This appendix is a brief review of a few useful algorithms based on the descent-direction method. We follow the notation usually used in the optimization literature. Suppose that the function to be minimized is the cost function $J(\mathbf{X})$ defined by Eq. 7. The descent-direction method is based on an iterative process which calculates a value of the control variable vector which minimizes the cost function. $\mathbf{X}_{k+1}$, the control variable vector at iteration $(k+1)$, has then to satisfy

$$J(\mathbf{X}_{k+1}) < J(\mathbf{X}_k). \tag{B1}$$

From an approximation $\mathbf{X}_k$ of the control variable vector at the iteration k, the new value $\mathbf{X}_{k+1}$ is computed in two stages: 1) find a descent-direction $\mathbf{d}_k$, and, 2) determine a positive step size α_k along $\mathbf{d}_k$. The second stage is called line search. $\mathbf{X}_{k+1}$ is then taken as

$$\mathbf{X}_{k+1} = \mathbf{X}_k + \alpha_k \mathbf{d}_k \tag{B2}$$

The way that the descent direction is computed will characterize the descent method. For instance, the simplest form for a descent direction is the negative value of the gradient of the function, *i.e.*,

$$\mathbf{d}_k = -\nabla J(\mathbf{X}_k) = -\mathbf{g}_k \tag{B3}$$

This method is called gradient or steepest descent method. As we shall see, most of the minimization techniques use a more general form of (B3), *i.e.*,

$$\mathbf{d}_k = -\mathbf{H}_k \mathbf{g}_k \tag{B4}$$

where $\mathbf{H}_k$ is a positive definite matrix.

The determination of the step size α_k requires some attention. Indeed, the step size has to be small enough in order to satisfy the relationship (B1). However, it has been shown that an arbitrarily small positive α_k can yield to the convergence to a non-stationary point. In most of the minimization algorithms, the dilemma is solved by choosing the step size such that it satisfies the Wolfe's conditions (Wolfe, 1969),

$$J(\mathbf{X}_{k+1}) \leq J(\mathbf{X}_k) + \gamma_1 \alpha_k \mathbf{g}_k^T \mathbf{d}_k, \tag{B5}$$

$$\mathbf{g}_{k+1}^T \mathbf{d}_k \geq \gamma_2 \mathbf{g}_k^T \mathbf{d}_k, \tag{B6}$$

where $0 < \gamma_1 < 0.5$ and $\gamma_1 < \gamma_2 < 1$, and T denotes the transpose. The first condition assures a sufficient decrease of the cost function while the second one avoids taking a step that is too small. In some cases, strong Wolfe's conditions have been found necessary. In that case, condition (B6) is replaced by

$$|\mathbf{g}_{k+1}^T \mathbf{d}_k| \leq |\gamma_2 \mathbf{g}_k^T \mathbf{d}_k|. \tag{B7}$$

An algorithm to find a point $\mathbf{X}_{k+1}$, referred to as Wolfe's point, which satisfies Eqs. B5 and B6 can be found in Lemaréchal (1981) while an algorithm for a point which satisfies the strong Wolfe's conditions can be found in Al-Baali and Fletcher (1986). Let us now examine different algorithms based on the descent method.

B.1 Conjugate-gradient method

In the conjugate-gradient method, the descent direction is computed by combining the gradient of the function with some information from the previous iterations. The general scheme of the conjugate-gradient algorithm is

1. take an initial value $\mathbf{X}_0$ and compute $\mathbf{d}_0 = -\mathbf{g}_0$

2. for k=0,1,2,...

- compute the step size α_k
- compute $\mathbf{X}_{k+1} = \mathbf{X}_k + \alpha_k \mathbf{d}_k$
- compute $\mathbf{g}_{k+1}$
- check if restart is necessary (see discussion below)
- if convergence (B1) then stop
- else compute $\mathbf{d}_{k+1} = -\mathbf{g}_{k+1} + \beta_{k+1}\mathbf{d}_k$

3. go to 2

β_{k+1} has been defined by Fletcher-Reeves as

$$\beta_{k+1} = \begin{cases} 0, & \text{for } k = 0, \\ \|\mathbf{g}_{k+1}\|^2 / \|\mathbf{g}_k\|^2, & \text{for } k = 1, 2, 3, ..., \end{cases} \tag{B8}$$

and by Polak-Riebere (Polak and Ribiere, 1969) as

$$\beta_{k+1} = \mathbf{g}_{k+1}^T (\mathbf{g}_{k+1} - \mathbf{g}_k) / \|\mathbf{g}_k\|^2, \quad \text{for } k \geq 0 \tag{B9}$$

Powell (1977, 1985) showed that Eq. B9 gives a much faster convergence than Eq. B8 and is preferred. It should also be mentioned that for minimization with regard to a large number of control variables, *e.g.*, like in oceanography, the conjugacy of the gradient might be lost after a few iterations and the direction of search might not be efficient. Powell (1977) suggested a criterion based on Beale (1972) for a restart to be performed, *i.e.*, when

$$|\mathbf{g}_{k+1}^T \mathbf{g}_k| \geq 0.2 \|\mathbf{g}_{k+1}\|^2 \tag{B10}$$

One has also to check if the direction of search is sufficiently downhill. Therefore, if the two following criteria are not satisfied a restart is indicated, *i.e.*,

$$\begin{aligned} \mathbf{g}_{k+1}^T \mathbf{d}_{k+1} &\leq -0.8 \|\mathbf{g}_{k+1}\|^2 \\ \mathbf{g}_{k+1}^T \mathbf{d}_{k+1} &\geq -1.2 \|\mathbf{g}_{k+1}\|^2 \end{aligned} \tag{B11}$$

The reader is referred to Derber (1985) and Long and Tacker (1989a,b) for computation of an optimal step-size, and to Beale (1972) and Navon and Legler (1987) for restart formulae.

The conjugate-gradient method has the advantage of being simple to implement but for large problems, the convergence can be slow and therefore costly if there is no preconditioning. One has to realize that at every iteration during the minimization process, the use of the cost function and its gradient requires a run of the model and its adjoint, respectively. The following techniques have a faster rate of convergence.

B.2 Newton method and its variations

B.2.1 Newton and truncated Newton methods

The Newton minimization method is based on the Newton method for solving the non-linear equations $\nabla J = 0$, *i.e.*, the equations are first linearized and then solved. In the minimization context, the linear equation system is

$$\nabla^2 J(\mathbf{X}_k)\mathbf{d}_k + \mathbf{g}_k = 0$$

where $\nabla^2 J(\mathbf{X}_k)$ is the Hessian matrix of the cost function at $\mathbf{X}_k$. This equation is solved for $\mathbf{d}_k$,

$$\mathbf{d}_k = -H_k^{-1}\mathbf{g}_k \tag{B12}$$

where $H_k^{-1} = [\nabla^2 J(\mathbf{X}_k)]^{-1}$ is the inverse of the Hessian matrix. This method requires the computation of the cost function, its first and second derivative at every point $\mathbf{X}_k$. While the convergence is faster than in the CG method, it can still be very costly, indeed impossible for large problems.

To overcome that difficulty, a truncated Newton method has been derived. This technique is based on the idea that the linear system (B12) can be partially solved using an iterative process. Dembo and Steibaug (1983), Schlick and Fogelson (1992a) considered a CG method to determine $\mathbf{d}_k$. In that case, the iteration is stopped

when the non-positiveness of the Hessian is detected. Another method, *i.e.*, Lanczos tridiagonalization, was used by O'Leary (1982). Two truncated Newton algorithms developed by Nash (1984a,b) and Schlick and Fogelson (1992a,b) are made publically available (see Wang *et al.*, 1992, 1995; Wang, 1993; Zou *et al.*, 1993a).

B.2.2 Quasi-Newton and limited memory quasi-Newton methods

As showed in the previous section, the Newton and truncated Newton methods require the computation of the Hessian. This computation can be very expensive. The quasi-Newton (QN), first developed by Davidon (1959), and limited-memory quasi-Newton (LMQN) methods circumvent this problem based on the idea that the Hessian matrix can be estimated as part of the iteration process using the information about the behavior of the cost function and its gradient to make rank-one or rank-two updates. These two methods do not actually compute the Hessian matrix but an approximation of it. The general scheme of the QN and LMQN methods are summarized as follows:

1. get an initial value $\mathbf{X}_0$ and $\mathbf{B}_0$
2. for k=0,1,2,...
   - $\mathbf{d}_k = -\mathbf{B}_k^{-1}\mathbf{g}_k$
   - get α_k satisfying the Wolfe's conditions (B5, B6)
   - compute $\mathbf{X}_{k+1} = \mathbf{X}_k + \alpha_k \mathbf{d}_k$
   - get $\mathbf{B}_{k+1}^{-1}$
3. if convergence (B1) then stop else go to 2

where $\mathbf{B}_k$ represents an approximation of the Hessian matrix at the kth iteration and $\mathbf{B}_k^{-1}$ its inverse. Often, the initial value for $\mathbf{B}_0$ is taken as the unit matrix or a

multiple of it. The updated matrix $\mathbf{B}_{k+1}$ has to satisfy the quasi-Newton condition, *i.e.*,

$$\mathbf{B}_{k+1}\mathbf{s}_k = \mathbf{y}_k$$

where $\mathbf{s}_k = \mathbf{X}_{k+1} - \mathbf{X}_k$ denotes the variation in the control variables during the kth iteration and $\mathbf{y}_k = \mathbf{g}_{k+1} - \mathbf{g}_k$ the change in the gradient.

The main difference between the QN and LMQN algorithms resides in the way of updating the matrix $\mathbf{B}_k$. The most effective algorithm to update $\mathbf{B}_k$ in the QN method is the BFGS (Broyden, Flecther, Golfard and Shanno) algorithm (Gilbert and Lemaréchal, 1989), *i.e.*,

$$\mathbf{B}_{k+1} = \mathbf{B}_k + \frac{\mathbf{y}_k\mathbf{y}_k^T}{\mathbf{y}_k^T\mathbf{s}_k} - \frac{\mathbf{B}_k\mathbf{s}_k\mathbf{s}_k^T\mathbf{B}_k}{\mathbf{s}_k^T\mathbf{B}_k\mathbf{s}_k} \tag{B13}$$

The inverse $\mathbf{B}_k^{-1}$ can be directly updated using the inverse BFGS formula, *i.e.*,

$$\mathbf{B}_{k+1}^{-1} = \left(I - \frac{\mathbf{s}_k\mathbf{y}_k^T}{\mathbf{y}_k^T\mathbf{s}_k}\right)\mathbf{B}_k^{-1}\left(I - \frac{\mathbf{y}_k\mathbf{s}_k^T}{\mathbf{y}_k^T\mathbf{s}_k}\right) + \frac{\mathbf{s}_k\mathbf{s}_k^T}{\mathbf{y}_k^T\mathbf{s}_k} \tag{B14}$$

From Eqs. B13 and B14, one can guess that the main disadvantage of the QN method is the need of a large storage capacity. While combining the low storage advantage of the CG method and the computational efficiency of the QN method, the LMQN method seems to be more appropriate for large problems. From Eq. B13, $\mathbf{B}_k$ is formed from $\mathbf{B}_0$ and k pairs $(\mathbf{y}_i, \mathbf{s}_i)$ with $0 \leq i < k$. Therefore, one only needs to store $\mathbf{B}_0$ (usually a multiple of the identity matrix) and the couples $(\mathbf{y}_i, \mathbf{s}_i)$ $(0 \leq i < k)$ in memory, and compute $\mathbf{B}_k^{-1}\mathbf{g}_k$ $(1 \leq i < k)$ with an appropriate algorithm. Unfortunately, the number of couples to store becomes quickly very large for large problems. The LMQN method, very often referred to as m-storage QN method, proposes to approximate $\mathbf{B}_k^{-1}$ at iteration k by computing its value from m couples and a starting matrix $\mathbf{B}_{0,k}^{-1}$, and only those m couples $(\mathbf{y}_i, \mathbf{s}_i)$ are stored. All the LMQN methods are based on that concept and only differ in the selection of the couples $(\mathbf{y}_i, \mathbf{s}_i)$, in the choice of the starting matrix $\mathbf{B}_{0,k}^{-1}$, in the method

for computing $\mathbf{B}_k^{-1}\mathbf{g}_k$ and also the presence or not of a restart. For example, the CONMIN algorithm of Shanno and Phua (1976) is a 2-storage LMQN method with a Beale restart, the algorithm from Buckley and LeNir (1983) is a m-storage LMQN method where m varies with the iteration index. For a description of these methods, see Zou *et al.* (1993a), Gilbert and Lemaréchal (1989).

The algorithm used in our study is based on Nocedal's proposal (1980) and has been implemented by Gilbert and Lemaréchal (1989). Nocedal's algorithm proposes to form $\mathbf{B}_k^{-1}$ by updating a matrix $\mathbf{B}_{0,k}^{-1}$ using the last m couples $(\mathbf{y}_i, \mathbf{s}_i)$ with $k-m \leq i < k-1$, and going from iteration k to $k+1$ the couple $(\mathbf{y}_{k-m}, \mathbf{s}_{k-m})$ is replaced by $(\mathbf{y}_k, \mathbf{s}_k)$. Therefore, $\mathbf{B}_k^{-1}$ is always updated with the latest information. In practice, the number of updates m is optimal between $3 \leq m \leq 7$. The step (2) in the schematic of the QN algorithm is then replaced by

- choose a positive definite matrix $\mathbf{B}_{0,k}^{-1}$
- for i=0 to i=m-1 compute $\mathbf{B}_{i,k}^{-1}$ using the inverse BFGS formula,

$$\mathbf{B}_{i+1,k}^{-1} = \left(I - \frac{\mathbf{s}_{k-m+i}\mathbf{y}_{k-m+i}^T}{\mathbf{y}_{k-m+i}^T\mathbf{s}_{k-m+i}}\right)\mathbf{B}_{i,k}^{-1}\left(I - \frac{\mathbf{y}_{k-m+i}\mathbf{s}_{k-m+i}^T}{\mathbf{y}_{k-m+i}^T\mathbf{s}_{k-m+i}}\right) + \frac{\mathbf{s}_{k-m+i}\mathbf{s}_{k-m+i}^T}{\mathbf{y}_{k-m+i}^T\mathbf{s}_{k-m+i}}$$

- take $\mathbf{B}_k^{-1} = \mathbf{B}_{m,k}^{-1}$

Since one has to choose $\mathbf{B}_{0,k}^{-1}$ at every iteration of the minimization, its choice is very important. The simplest choice would be to take a diagonal matrix multiple of the identity matrix $\mathbf{I}$, *i.e.*,

$$\mathbf{B}_{0,k}^{-1} = \delta_{k-1}\mathbf{I} \quad \text{where} \quad \delta_{k-1} = \frac{\mathbf{y}_{k-1}^T\mathbf{s}_{k-1}}{|\mathbf{y}_{k-1}|^2}$$

In the algorithm proposed by Gilbert and Lemaréchal (1989), the matrix $\mathbf{B}_{0,k}^{-1}$ is taken as a diagonal matrix $\mathbf{D}_k$ which is updated at each iteration k from $\mathbf{D}_{k-1}$. When $k = 1$, the diagonal matrix is taken as

$$\mathbf{D}_1 = \frac{\mathbf{y}_0^T\mathbf{s}_0}{\|\mathbf{y}_0\|^2}I$$

When $k > 1$, each element j of the matrix is updated as follows

$$\mathbf{D}_{k+1}^{(j)} = \left(\frac{\langle \mathbf{D}_k \mathbf{y}_k, \mathbf{y}_k \rangle}{\langle \mathbf{y}_k, \mathbf{s}_k \rangle \mathbf{D}_k^{(j)}} + \frac{\langle \mathbf{y}_k, \mathbf{e}_j \rangle^2}{\langle \mathbf{y}_k, \mathbf{s}_k \rangle} - \frac{\langle \mathbf{D}_k \mathbf{y}_k, \mathbf{y}_k \rangle \langle \mathbf{s}_k, \mathbf{e}_j \rangle^2}{\langle \mathbf{y}_k, \mathbf{s}_k \rangle \langle \mathbf{D}_k^{-1} \mathbf{s}_k, \mathbf{s}_k \rangle (\mathbf{D}_k^{(j)})^2} \right)^{-1}.$$

where $\mathbf{e}_j$ is a vector with its jth element equal to 1 and the other elements equal to zero, and $\langle ., . \rangle$ represents the inner product.

This algorithm routine and its double precision version are called M1QN3 and N1QN3 in the French optimization library MODULOPT. It has been shown to be the LMQN method with the fastest convergence rate (Zou *et al.*, 1993; Gilbert and Lemaréchal, 1989, Liu and Nocedal, 1989). Note that the Nocedal (1980) LMQN is practically identical to M1QN3.

C Performance of the tidal data assimilation

This appendix gives the root-mean square error (rms), the relative average error (E) and correlation coefficient between modeled and predicted elevations for November 2 to November 19, 1983. The various quantities have been computed for each day of the assimilation experiment using Eqs. 22, 23, 24. In each table, the first six stations are the permanent tide gauge stations and the last six stations are the comparison stations. Their locations are plotted in Fig. 12. For each station, the first number corresponds to the case when modeled elevations were obtained using a constant bottom drag coefficient equal to 0.002. The second number corresponds to the case when the modeled elevations were computed using the estimated bottom drag coefficient from the data assimilation.

Station Name	November 1983								
	2	3	4	5	6	7	8	9	10
Havre de Grace, MD	8.4	9.4	10.6	11.4	11.8	11.6	10.9	9.8	8.4
	8.2	8.5	8.8	9.1	8.9	8.8	8.5	8.2	7.7
Baltimore, MD	6.4	6.8	7.2	7.5	7.6	7.4	7.1	6.3	5.6
	4.4	4.6	4.6	4.4	4.1	3.5	3.0	2.5	2.2
Annapolis, MD	4.5	5.1	5.6	6.1	6.4	6.3	6.0	5.3	4.5
	4.1	4.4	4.6	4.5	4.2	3.6	3.1	2.5	1.9
Cambridge, MD	8.5	9.0	9.5	9.8	9.8	9.5	8.9	8.2	7.3
	3.7	3.8	3.9	3.9	3.9	4.0	4.0	3.9	3.9
Solomons Is., MD	6.4	7.0	7.3	7.4	7.2	6.6	5.8	5.0	4.1
	4.2	4.6	4.8	4.9	4.6	4.0	3.4	2.7	2.2
Lewisetta, MD	5.6	6.2	6.6	6.7	6.5	5.9	5.1	4.2	3.4
	4.5	4.8	5.1	5.1	4.9	4.4	3.7	3.0	2.3
Gloucester Pt, VA	10.8	11.8	12.4	12.5	12.0	11.0	9.5	7.8	6.1
	6.6	7.0	7.3	7.2	6.9	6.2	5.3	4.2	3.1
Kiptopeake, VA	6.7	7.5	8.1	8.4	8.2	7.5	6.5	5.4	4.4
	3.4	3.7	4.0	4.1	4.0	3.7	3.4	3.1	3.0
Hampton Roads, VA	9.3	10.2	10.8	11.0	10.7	9.8	8.5	7.1	5.6
	6.2	6.8	7.3	7.4	7.2	6.7	5.8	4.9	3.9
CBBT, VA	4.6	5.2	5.6	5.8	5.6	5.1	4.4	3.6	2.9
	3.5	3.9	4.2	4.4	4.3	4.1	3.6	3.2	3.0
Betterton, MD	7.8	8.7	9.7	10.5	10.9	10.8	10.2	9.2	7.9
	5.6	6.0	6.3	6.4	6.3	5.9	5.4	4.9	4.3
Matapeake, MD	4.8	5.3	5.8	6.2	6.4	6.2	5.8	5.1	4.2
	3.8	4.2	4.3	4.3	3.9	3.5	3.0	2.6	2.3
Avalon, MD	7.0	7.9	8.4	8.6	8.2	7.5	6.6	5.4	4.3
	4.3	4.9	5.3	5.4	5.1	4.5	3.8	3.0	2.4
Chesapeake Bch, MD	4.3	4.8	5.2	5.5	5.5	5.2	4.8	4.1	3.4
	4.4	4.7	4.9	4.9	4.7	4.3	3.8	3.2	2.7
Colonial Bch, VA	10.6	11.4	12.0	12.2	11.8	10.9	9.8	8.5	7.1
	8.4	9.1	9.5	9.6	9.2	8.5	7.6	6.5	5.5
Holland Bar Lt, MD	6.1	6.6	6.8	6.7	6.2	5.5	4.6	3.7	3.0
	4.1	4.4	4.6	4.7	4.5	4.0	3.5	2.9	2.4
Guardshore, VA	9.1	9.9	10.4	10.5	10.1	9.2	8.0	6.8	5.6
	5.3	5.5	5.8	6.2	6.6	6.6	6.3	5.8	5.3
Rappahannock, VA	5.1	5.7	6.1	6.2	6.0	5.4	4.8	4.1	3.5
	3.7	4.0	4.2	4.3	4.2	3.9	3.7	3.5	3.4
New Pt Comf Sh, VA	7.4	7.8	8.1	8.2	8.0	7.5	6.7	5.8	5.0
	5.1	5.2	5.2	5.1	4.8	4.4	4.0	3.6	3.4

Table 6: Root-mean square error (cm) at ten permanent and nine comparison tide gauge stations for November 2 to November 10, 1983. The first number corresponds to the experiment with $c_D = 0.002$ and the second number corresponds to the recovery experiment.

Station Name	November 1983								
	11	12	13	14	15	16	17	18	19
Havre de Grace, MD	6.9	5.5	4.3	3.3	2.5	2.4	3.1	4.7	6.8
	6.9	6.0	4.9	4.0	3.2	2.8	3.1	4.0	5.0
Baltimore, MD	4.9	4.3	3.9	3.6	3.3	3.3	3.7	4.4	5.4
	2.1	2.3	2.4	2.6	2.7	2.6	2.5	2.6	2.9
Annapolis, MD	3.7	2.9	2.4	1.9	1.7	1.9	2.4	3.3	4.3
	1.5	1.2	1.1	1.2	1.3	1.5	1.8	2.3	2.9
Cambridge, MD	6.4	5.6	5.1	4.7	4.5	4.5	4.9	5.8	6.9
	4.0	4.0	3.9	3.8	3.7	3.5	3.4	3.4	3.7
Solomons Is, MD	3.4	2.9	2.6	2.5	2.6	2.9	3.4	3.9	4.7
	2.0	2.0	2.1	2.3	2.4	2.6	2.9	3.2	3.6
Lewisetta, MD	2.7	2.2	2.0	1.9	2.1	2.4	2.8	3.4	4.1
	1.6	1.3	1.2	1.3	1.6	1.8	2.1	2.6	3.2
Gloucester Pt, VA	4.6	3.4	2.7	2.6	2.9	3.5	4.5	5.8	7.3
	2.3	1.8	1.6	1.8	2.1	2.5	3.0	3.8	4.8
Kiptopeake, VA	3.8	3.5	3.5	3.6	3.6	3.5	3.4	3.4	3.7
	3.1	3.4	3.6	3.6	3.6	3.5	3.3	3.0	2.9
Hampton Roads, VA	4.3	3.3	2.9	2.9	3.0	3.3	3.9	4.7	5.9
	3.2	2.8	2.7	2.8	2.9	3.0	3.3	3.8	4.4
CBBT, VA	2.6	2.6	2.8	3.0	3.0	2.8	2.6	2.3	2.5
	2.9	3.1	3.3	3.4	3.3	3.3	3.1	3.1	3.2
Betterton, MD	6.5	5.1	3.8	2.8	2.4	2.8	3.9	5.4	7.0
	3.8	3.2	2.7	2.2	2.1	2.4	3.0	3.7	4.5
Matapeake, MD	3.4	2.8	2.4	2.1	1.9	1.9	2.4	3.3	4.3
	2.0	1.7	1.5	1.5	1.6	1.7	1.9	2.3	2.8
Avalon, MD	3.5	3.2	3.3	3.5	3.7	3.8	4.1	4.5	5.1
	2.1	2.2	2.5	2.9	3.1	3.2	3.1	2.9	3.1
Chesapeake Bch, MD	2.7	2.2	1.9	1.7	1.6	1.6	1.9	2.5	3.4
	2.2	1.8	1.6	1.6	1.7	1.9	2.2	2.6	3.2
Colonial Bch, VA	5.7	4.6	3.8	3.5	3.6	4.3	5.4	6.7	8.0
	4.5	3.6	3.1	2.9	3.0	3.4	4.1	5.0	6.0
Holland Bar Lt, MD	2.5	2.4	2.4	2.6	2.8	3.1	3.6	4.1	4.8
	2.1	2.0	2.1	2.3	2.4	2.6	2.8	3.1	3.5
Guardshore, VA	4.8	4.3	3.9	3.8	3.8	4.0	4.5	5.3	6.6
	4.7	3.9	3.3	3.1	3.0	3.1	3.3	3.7	4.3
Rappahannock, VA	3.2	3.1	3.1	3.0	3.0	3.0	3.0	3.1	3.3
	3.3	3.3	3.2	3.1	3.1	3.2	3.3	3.6	4.0
New Pt Comf Sh, VA	4.3	3.7	3.2	2.8	2.6	2.6	2.8	3.2	3.9
	3.2	3.1	2.8	2.6	2.3	2.3	2.5	2.9	3.5

Table 7: Root-mean square error (cm) at ten permanent and nine comparison tide gauge stations for November 11 to November 19, 1983. The first number corresponds to the experiment with $c_D = 0.002$ and the second number corresponds to the recovery experiment.

Station Name	November 1983								
	2	3	4	5	6	7	8	9	10
Havre de Grace, MD	10.93	12.54	14.50	16.12	16.90	16.69	15.66	14.01	11.94
	6.15	6.08	5.97	5.73	5.40	5.27	5.29	5.43	5.62
Baltimore, MD	17.78	18.63	19.53	20.34	20.74	20.48	19.58	18.12	16.33
	5.81	5.69	5.26	4.58	3.83	3.08	2.37	1.84	1.67
Annapolis, MD	12.05	13.54	15.27	16.68	17.56	17.73	17.00	15.26	12.83
	6.62	7.03	6.78	6.00	4.91	3.86	2.96	2.18	1.56
Cambridge, MD	14.66	15.47	16.32	16.90	17.16	17.16	16.87	16.18	15.28
	2.07	1.94	1.88	1.85	1.90	2.06	2.30	2.67	3.32
Solomons Is, MD	14.05	15.26	16.09	16.30	15.92	14.93	13.32	11.50	9.84
	4.49	4.79	4.95	4.91	4.52	3.87	3.18	2.53	2.09
Lewisetta, MD	9.08	10.23	11.04	11.32	10.99	10.09	8.71	7.07	5.51
	4.13	4.34	4.50	4.55	4.38	3.93	3.29	2.57	1.84
Gloucester Pt, VA	8.93	9.68	10.25	10.44	10.12	9.32	8.13	6.70	5.18
	2.78	2.85	2.92	2.89	2.74	2.47	2.07	1.58	1.14
Kiptopeake, VA	2.44	2.86	3.20	3.41	3.40	3.18	2.80	2.35	1.98
	0.55	0.61	0.67	0.71	0.71	0.68	0.65	0.66	0.80
Hampton Roads, VA	5.96	6.59	7.10	7.36	7.24	6.72	5.89	4.88	3.85
	2.39	2.60	2.80	2.92	2.91	2.75	2.44	2.07	1.73
CBBT, VA	1.08	1.27	1.42	1.51	1.51	1.40	1.21	0.99	0.83
	0.60	0.67	0.75	0.81	0.84	0.83	0.79	0.76	0.83
Betterton, MD	12.71	14.58	16.80	18.67	19.65	19.50	18.23	16.16	13.60
	3.87	4.04	4.01	3.91	3.66	3.28	2.88	2.59	2.40
Matapeake, MD	9.96	11.25	12.61	13.62	14.15	14.06	13.20	11.57	9.53
	4.31	4.57	4.47	4.06	3.44	2.81	2.30	1.96	1.77
Avalon, MD	11.34	12.95	14.06	14.34	13.72	12.40	10.64	8.59	6.61
	2.93	3.39	3.67	3.67	3.34	2.83	2.27	1.73	1.36
Chesapeake Bch, MD	8.64	9.63	10.67	11.34	11.45	11.01	10.08	8.70	7.05
	5.87	6.09	6.05	5.80	5.29	4.59	3.88	3.27	2.81
Colonial Bch, VA	21.33	22.88	23.94	24.29	23.64	22.10	20.07	17.61	14.94
	9.15	9.58	9.83	9.86	9.50	8.83	8.03	7.18	6.30
Holland Bar Lt, MD	7.94	8.54	8.81	8.60	7.91	6.85	5.60	4.35	3.43
	2.58	2.74	2.86	2.91	2.82	2.58	2.28	2.00	1.76
Guardshore, VA	6.89	7.52	7.98	8.06	7.77	7.14	6.28	5.37	4.61
	1.71	1.71	1.78	1.99	2.33	2.60	2.79	2.94	3.13
Rappahannock, VA	4.32	4.93	5.40	5.57	5.41	4.98	4.41	3.93	3.74
	1.83	1.93	2.02	2.06	2.04	2.02	2.08	2.32	2.82
New Pt Comf Sh, VA	4.94	5.14	5.29	5.36	5.28	5.01	4.58	4.12	3.72
	2.00	1.94	1.83	1.72	1.59	1.48	1.40	1.39	1.52

Table 8: Relative average error (%) at ten permanent and nine comparison tide gauge stations for November 2 to November 10, 1983. The first number corresponds to the experiment with $c_D = 0.002$ and the second number corresponds to the recovery experiment.

Station Name	November 1983								
	11	12	13	14	15	16	17	18	19
Havre de Grace, MD	9.64	7.37	5.30	3.50	2.16	1.73	2.56	4.86	8.13
	5.73	5.65	5.29	4.40	3.03	2.18	2.16	2.62	3.10
Baltimore, MD	14.64	13.62	13.31	13.03	12.02	10.76	10.66	12.08	14.43
	1.95	2.72	3.96	5.54	6.47	5.58	4.04	3.18	3.04
Annapolis, MD	10.30	8.14	6.50	5.35	4.54	4.65	6.13	8.73	11.68
	1.12	0.93	1.07	1.62	2.31	2.73	2.96	3.30	3.83
Cambridge, MD	14.35	13.47	12.53	11.32	9.94	8.94	9.04	10.28	12.08
	4.28	5.44	6.39	6.71	6.16	5.01	3.77	2.98	2.73
Solomons Is, MD	8.49	7.54	7.09	7.01	7.04	7.14	7.60	8.52	9.80
	2.12	2.75	3.90	4.95	5.38	5.20	4.74	4.43	4.46
Lewisetta, MD	4.32	3.58	3.30	3.35	3.58	3.93	4.42	5.14	6.07
	1.26	0.96	1.01	1.35	1.75	1.95	2.03	2.29	2.81
Gloucester Pt, VA	3.77	2.65	1.97	1.79	1.97	2.41	3.12	4.06	5.18
	0.79	0.62	0.65	0.83	1.00	1.12	1.27	1.54	1.92
Kiptopeake, VA	1.83	1.99	2.37	2.57	2.30	1.76	1.28	1.01	0.97
	1.15	1.72	2.30	2.51	2.20	1.64	1.10	0.73	0.54
Hampton Roads, VA	2.93	2.31	2.06	2.04	2.03	2.01	2.14	2.48	3.04
	1.51	1.51	1.69	1.82	1.75	1.58	1.45	1.44	1.58
CBBT, VA	0.86	1.13	1.58	1.85	1.64	1.15	0.70	0.45	0.40
	1.06	1.52	2.05	2.25	1.96	1.46	1.02	0.74	0.64
Betterton, MD	10.82	8.06	5.51	3.58	2.78	3.42	5.47	8.52	11.92
	2.25	2.11	1.94	1.76	1.72	2.02	2.54	3.00	3.32
Matapeake, MD	7.62	6.25	5.45	4.88	4.18	3.86	4.69	6.62	9.06
	1.71	1.71	1.80	2.13	2.50	2.46	2.32	2.46	2.87
Avalon, MD	5.29	5.21	6.23	7.29	7.48	7.02	6.62	6.78	7.48
	1.37	1.97	3.18	4.42	4.95	4.42	3.26	2.31	2.03
Chesapeake Bch, MD	5.50	4.43	3.94	3.68	3.24	2.80	3.07	4.34	6.28
	2.51	2.35	2.42	2.79	3.25	3.46	3.50	3.69	4.21
Colonial Bch, VA	12.22	9.64	7.62	6.59	6.74	8.04	10.23	12.91	15.58
	5.40	4.59	4.02	3.77	3.88	4.23	4.79	5.59	6.45
Holland Bar Lt, MD	3.09	3.35	3.93	4.46	4.71	4.80	4.99	5.36	5.93
	1.72	2.00	2.53	3.05	3.24	2.98	2.57	2.36	2.42
Guardshore, VA	4.15	3.98	3.89	3.79	3.58	3.32	3.27	3.65	4.47
	3.14	2.78	2.35	2.18	2.02	1.74	1.47	1.41	1.50
Rappahannock, VA	3.97	4.61	5.31	5.47	4.88	3.93	3.10	2.61	2.50
	3.59	4.47	5.05	5.07	4.66	4.00	3.38	3.00	2.86
New Pt Comf Sh, VA	3.43	3.17	2.78	2.24	1.73	1.44	1.38	1.50	1.74
	1.75	2.01	2.07	1.79	1.36	1.09	1.03	1.11	1.29

Table 9: Relative average error (%) at ten permanent and nine comparison tide gauge stations for November 11 to November 19, 1983. The first number corresponds to the experiment with $c_D = 0.002$ and the second number corresponds to the recovery experiment.

Station Name	November 1983								
	2	3	4	5	6	7	8	9	10
Havre de Grace, MD	0.980	0.982	0.981	0.978	0.973	0.968	0.964	0.961	0.957
	0.967	0.964	0.963	0.966	0.970	0.971	0.971	0.970	0.969
Baltimore, MD	0.972	0.975	0.976	0.977	0.977	0.975	0.974	0.972	0.969
	0.966	0.968	0.972	0.977	0.982	0.985	0.988	0.989	0.988
Annapolis, MD	0.985	0.984	0.979	0.974	0.970	0.969	0.969	0.971	0.973
	0.981	0.979	0.980	0.982	0.986	0.989	0.991	0.993	0.994
Cambridge, MD	0.995	0.996	0.996	0.994	0.990	0.986	0.980	0.973	0.966
	0.993	0.994	0.995	0.994	0.992	0.988	0.984	0.980	0.973
Solomons Is, MD	0.988	0.987	0.987	0.986	0.985	0.983	0.981	0.979	0.976
	0.992	0.991	0.990	0.990	0.990	0.992	0.992	0.992	0.991
Lewisetta, MD	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.996	0.995
	0.991	0.990	0.989	0.988	0.988	0.989	0.991	0.993	0.995
Gloucester Pt, VA	0.972	0.971	0.971	0.971	0.973	0.976	0.980	0.984	0.988
	0.989	0.989	0.990	0.990	0.991	0.992	0.993	0.995	0.997
Kiptopeake, VA	0.998	0.998	0.998	0.999	0.999	0.999	0.999	0.998	0.996
	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.995
Hampton Roads, VA	0.991	0.991	0.991	0.992	0.992	0.994	0.996	0.997	0.998
	0.998	0.999	0.999	0.999	0.998	0.998	0.998	0.998	0.997
CBBT, VA	0.998	0.998	0.998	0.998	0.999	0.999	0.999	0.999	0.998
	1.000	1.000	0.999	0.999	0.999	0.999	0.999	0.998	0.995
Betterton, MD	0.996	0.994	0.990	0.985	0.980	0.977	0.975	0.975	0.975
	0.994	0.993	0.993	0.992	0.991	0.991	0.992	0.992	0.993
Matapeake, MD	0.995	0.995	0.992	0.987	0.982	0.978	0.975	0.974	0.972
	0.994	0.993	0.993	0.994	0.995	0.996	0.996	0.996	0.995
Avalon, MD	0.993	0.991	0.989	0.988	0.988	0.990	0.992	0.993	0.993
	0.992	0.990	0.987	0.986	0.987	0.990	0.993	0.996	0.998
Chesapeake Bch, MD	0.997	0.997	0.996	0.992	0.989	0.986	0.984	0.981	0.979
	0.996	0.997	0.997	0.997	0.997	0.998	0.997	0.997	0.995
Colonial Bch, VA	0.927	0.925	0.924	0.924	0.927	0.932	0.939	0.947	0.955
	0.917	0.913	0.911	0.910	0.913	0.919	0.925	0.933	0.941
Holland Bar Lt, MD	0.992	0.993	0.994	0.995	0.995	0.995	0.994	0.992	0.989
	0.997	0.996	0.996	0.996	0.997	0.997	0.998	0.998	0.997
Guardshore, VA	0.991	0.992	0.991	0.991	0.990	0.989	0.987	0.984	0.980
	0.990	0.990	0.990	0.988	0.985	0.982	0.979	0.977	0.974
Rappahannock, VA	0.995	0.995	0.994	0.994	0.992	0.990	0.987	0.983	0.976
	0.998	0.998	0.997	0.997	0.996	0.995	0.992	0.988	0.982
New Pt Comf Sh, VA	0.979	0.980	0.982	0.985	0.988	0.992	0.995	0.997	0.997
	0.991	0.992	0.993	0.995	0.997	0.998	0.998	0.998	0.996

Table 10: Correlation coefficient at ten permanent and nine comparison tide gauge stations for November 2 to November 10, 1983. The first number corresponds to the experiment with $c_D = 0.002$ and the second number corresponds to the recovery experiment.

Station Name	November 1983								
	11	12	13	14	15	16	17	18	19
Havre de Grace, MD	0.955	0.955	0.959	0.969	0.982	0.989	0.990	0.987	0.981
	0.967	0.962	0.958	0.962	0.975	0.984	0.986	0.983	0.980
Baltimore, MD	0.964	0.956	0.945	0.936	0.939	0.954	0.968	0.973	0.974
	0.984	0.976	0.966	0.956	0.951	0.960	0.974	0.984	0.988
Annapolis, MD	0.975	0.975	0.973	0.973	0.977	0.981	0.980	0.975	0.971
	0.995	0.996	0.995	0.993	0.991	0.991	0.992	0.992	0.990
Cambridge, MD	0.958	0.950	0.945	0.947	0.956	0.968	0.977	0.981	0.982
	0.965	0.958	0.956	0.960	0.966	0.974	0.982	0.987	0.988
Solomons Is, MD	0.972	0.967	0.963	0.961	0.964	0.968	0.973	0.976	0.978
	0.988	0.980	0.969	0.962	0.962	0.967	0.973	0.980	0.985
Lewisetta, MD	0.993	0.990	0.987	0.985	0.985	0.987	0.989	0.991	0.993
	0.997	0.998	0.996	0.993	0.992	0.993	0.995	0.997	0.998
Gloucester Pt, VA	0.992	0.994	0.996	0.997	0.997	0.997	0.995	0.991	0.985
	0.998	0.998	0.998	0.998	0.998	0.999	0.998	0.996	0.993
Kiptopeake, VA	0.991	0.985	0.979	0.976	0.979	0.986	0.992	0.996	0.998
	0.990	0.984	0.978	0.976	0.979	0.986	0.991	0.996	0.998
Hampton Roads, VA	0.997	0.995	0.992	0.990	0.991	0.995	0.997	0.998	0.997
	0.995	0.992	0.989	0.988	0.990	0.993	0.996	0.998	0.999
CBBT, VA	0.995	0.990	0.985	0.983	0.985	0.991	0.995	0.998	0.999
	0.992	0.987	0.982	0.980	0.983	0.989	0.994	0.997	0.999
Betterton, MD	0.976	0.978	0.981	0.986	0.990	0.990	0.983	0.976	0.970
	0.992	0.990	0.988	0.988	0.989	0.988	0.986	0.984	0.983
Matapeake, MD	0.970	0.966	0.962	0.962	0.970	0.979	0.981	0.980	0.977
	0.993	0.991	0.989	0.987	0.986	0.990	0.994	0.995	0.995
Avalon, MD	0.988	0.977	0.963	0.955	0.958	0.969	0.981	0.990	0.994
	0.994	0.984	0.971	0.964	0.965	0.971	0.982	0.991	0.996
Chesapeake Bch, MD	0.977	0.974	0.971	0.972	0.979	0.987	0.991	0.990	0.987
	0.993	0.990	0.986	0.984	0.984	0.987	0.992	0.995	0.996
Colonial Bch, VA	0.961	0.967	0.973	0.976	0.976	0.973	0.967	0.960	0.954
	0.949	0.956	0.963	0.967	0.968	0.966	0.960	0.952	0.945
Holland Bar Lt, MD	0.984	0.977	0.970	0.966	0.968	0.973	0.979	0.984	0.987
	0.994	0.988	0.980	0.974	0.974	0.979	0.985	0.990	0.993
Guardshore, VA	0.975	0.971	0.968	0.967	0.969	0.973	0.979	0.985	0.989
	0.973	0.975	0.979	0.980	0.982	0.986	0.989	0.991	0.991
Rappahannock, VA	0.967	0.958	0.950	0.950	0.957	0.968	0.977	0.984	0.988
	0.973	0.964	0.958	0.958	0.965	0.974	0.984	0.990	0.994
New Pt Comf Sh, VA	0.996	0.994	0.992	0.993	0.994	0.996	0.996	0.995	0.994
	0.993	0.989	0.988	0.990	0.993	0.995	0.996	0.996	0.996

Table 11: Correlation coefficient at ten permanent and nine comparison tide gauge stations for November 11 to November 19, 1983. The first number corresponds to the experiment with $c_D = 0.002$ and the second number corresponds to the recovery experiment.

Curriculum Vitae

Yvette Huberte Spitz

Date of Birth : July 16, 1960
Place of Birth: Heure-le-Romain, Belgium

Education

M.Sc., Oceanography, 1990, Florida State University, Tallahassee, Florida.
Licence, Oceanology, 1985, Liège University, Liège, Belgium.
Licence, Physics, 1983, Liège University, Liège, Belgium.

Reviewed Publications

Jamart, B.M., J. Ozer, and Y. Spitz, 1986: Bottom stress and free oscillations . *Advanced Physical Oceanographic Numerical Modelling*, J.J. O'Brien, Ed., D. Reidel Publishing Company, 581–598.
Spitz, Y.H. and D. Nof, 1991: Separation of boundary currents due to bottom topography. *Deep-Sea Research*, Vol. 38, **1**, 1–20.
Lawson, L.M., Y.H. Spitz, E.E. Hofmann and R.B. Long, 1995: A data assimilation technique applied to a predator-prey model. *Bullletin of Mathematical Biology*, Vol. 57, **4**, 593–617.
Lawson, L. M., E. E. Hofmann, and Y. H. Spitz, 1995: Time series sampling and data assimilation in a simple marine ecosystem model. (submitted).

Other Publications and Abstracts

Demuth, C., and Y. Spitz, 1985: Modèle mathématique de dispersion des radionucléides déversés dans le Nord Atlantique. *Final report to ONDRAF(Organisme National des Déchets Radioactifs et des Matières Fissiles)*, 169 pp.
Ozer, J., B.M. Jamart and Y. Spitz, 1986: Preliminary description and validation of the "FU-model", a semi-implicit 3-D model with continuous representation of the velocity profiles. *MUMM's contribution to BSEX Task IV, Technical Report No. 2*, 52 pp.
Jamart, B.M., Y. Spitz, and J. Ozer, 1986: Digitization of the coastlines and bathymetry of the North Sea and surrounding waters. *MUMM's contribution to BSEX Task III, Technical Report No. 2*, 41 pp + 3 appendices.
Spitz, Y., B.M. Jamart, and J. Ozer, 1987: Numerical experiments with a vertical plane, 2.5-D hydrodynamical model. *MUMM's contribution to BSEX Task IV, Technical Report No. TR05*, 106 pp.
Ozer, J., B.M. Jamart, and Y. Spitz, 1988: Wind-driven flows in an enclosed basin, a comparison of numerical models. *MUMM's contribution to BSEX, Technical Report No. TR06*, 140 pp + 1 appendix.
Spitz, Y.H., and D. Nof. Effect of Bottom Topography on a Coastal Current. *EOS, Transactions American Geophysical Union*, **71(2)**, January 1990.
Spitz, Y.H., and J.M. Klinck. Recovery of Circulation and Forcing in an Enclosed Basin by Assimilation of Tide Gauge Observations. *EOS, Transactions American Geophysical Union*, **75(3)**, January 1994.

Lawson, L.M., Y.H. Spitz, and E.E. Hofmann. Data assimilation techniques applied to a marine ecosystem model. *EOS, Transactions American Geophysical Union*, **75(3)**, January 1994.

Professional Affiliations

Research Assistant, GeoHydrodynamics and Environmental Research, Liège University, Liège, Belgium, 1983–1984.
Assistant Oceanographer, Management Unit of the Mathematical Models of the North Sea and the Scheldt Estuary, Brussels, Belgium, 1984-1987.
Graduate Research Assistant, Florida State University, Tallahassee, Florida, 1987-1991.
Graduate Research Assistant, Old Dominion University, Norfolk, Virginia, 1991-present.
Student, International Summer School on Assimilation of Meteorological and Oceanographical Observations, La Garde, France, August 1993.

Award

Phi Kappa Phi, May 1993

Memberships

American Meteorological Society
The Oceanography Society
American Geophysical Union