Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Continental Shelf Research 26 (2006) 539–553

CONTINENTAL SHELF RESEARCH

www.elsevier.com/locate/csr

A data assimilation procedure for operational prediction of storm surge in the northern Adriatic Sea

P. Lionello[a,*], A. Sanna[a,1], E. Elvini[b], R. Mufato[b,2]

[a]Department of Material Science, University of Lecce, via per Arnesano, 73100 Lecce, Italy
[b]University of Padova, Italy

Received 27 January 2005; received in revised form 26 December 2005; accepted 11 January 2006

Abstract

A procedure designed for the operational prediction of storm surge in northern Adriatic Sea is presented. Its main purpose is to provide an early warning for the flood of the Venice city centre. The procedure uses a hydro-dynamical shallow water model (forced by the wind stress and sea level pressure fields) and assimilates hourly sea level observations (available at the "Aqua Alta" research platform, 15 km offshore the Venetian littoral) in order to produce an optimal surge forecast. A cost function describing the discrepancy between model results and observations is defined and the adjoint and the conjugate gradient methods are used for the computation of the cost function gradient and the search of its minimum, respectively. Each operational simulation is split into an analysis period (whose optimal length has been found to be 3 days) and a forecast period (prediction has been considered up to the 3-day range). During the analysis, the observations are assimilated in the model in order to identify the optimal initial condition for the surge prediction carried out for the following forecast period. A penalty term, which improves the stability of the assimilation procedure, has been identified and included in the cost function. The results show that the procedure is capable of compensating for the errors (due to both inaccurate meteorological forcing and model shortcomings) and effectively improves the reliability of the storm surge forecast. Moreover, this study proves that the main source of error for the short and medium range forecast, in such a weakly dissipative system, resides in an inaccurate initial condition, which can be improved by a variational data assimilation procedure.

Keywords: Storm surge; Data assimilation; Adjoint model; Operational; Prediction

1. Introduction

Storm surge in the northern Adriatic Sea represents a constant threat to its northern coastal areas, which consist of the borders of the Po Valley and the Friuli-Venetian planes. Both are low level and flat regions, with lagoons closed by sandy narrow strips of land and areas below the sea level (Bondesan et al., 1995).

The most critical situation is represented by the city and the lagoon of Venice with its monumental and environmental heritage, which on 4th November 1966 was hit by the highest surge ever recorded. The sea level raised about 180 cm above the mean

*Corresponding author. Tel.: +39 049 8277289; fax: +39 049 8277282.
E-mail address: piero.lionello@unile.it (P. Lionello).
[1]Present affiliation: ARPA-Piemonte.
[2]Present affiliation: ARPA-Veneto.

0278-4343/$ - see front matter
doi:10.1016/j.csr.2006.01.003

sea level (MSL) and persisted for more than 15 h above the 100 cm level (Canestrelli et al., 2001), which at that time corresponded to the critical value above which more than two thirds of Venice were flooded (Frassetto, 1976). Electric power supply and telephone connections were cut off in Venice and in the islands of the lagoon. High waves swept the coastal defences, made a breach about 4 km wide through which the lagoon was temporarily joined to the Adriatic Sea. Though it was necessary to evacuate with motorboats about 4000 people from two coastal villages (Pellestrina and S. Pietro in Volta) there were, fortunately, only three casualties caused by storm surge. After the storm, it took more than one week to repair the phone line and re-install the electric power supply. Economic loss were evaluated about 40 billions of Italian lire (nowadays equivalent to about 400 million euros).

The November 1966 surge level was clearly exceptional and return time is likely larger than 250 years (Lionello, 2005). Moreover, in 1966 the city structure was not prepared to face such event. Afterwards, power supply and telephone lines have been completely renovated and restructured, coastal defences have been rebuilt and modified, the street level of most of the city has been (or is going to be) brought to 1 m above the MSL, though the situation of the oldest central monumental area, whose level is between 60 and 70 cm above the MSL, remains critical. Therefore, nowadays, an event comparable to that of November 1966 would produce much lower damage and could be sustained much better. However, in a long time perspective, that is on decadal to centennial time scale, sea level rise (estimated from 9 to 88 cm in the 21st century, IPCC 3rd report, 2001) and natural subsidence, which is associated to the subduction of the Adriatic plate under the Apennines, estimated about 1 mm per year (Carminati et al., 2005), are going to increase the frequency and the level of the storm surge events. Moreover, the planned flood defence system is based on mobile barriers, which will operate closing the three inlets of the Venetian lagoon for surges predicted to be higher than 80 cm and will require a forecast system providing a precise warning 6 h in advance (Cecconi, 1999). A reliable forecast system is an important component of the long time strategy to control the recurrent floods.

The Venice City Council has established a dedicated centre, called CPSM (Centro Previsioni e Segnalazioni Maree, Centre for Tide Prediction and Warning), which provides regular forecast of the sea level. The main forecast system of CPSM is based on a linear regressive model (Tomasin, 1972; Canestrelli, 1999), producing the surge forecast on the basis of the atmospheric surface pressure at a set of stations surrounding the Adriatic Sea and of the observed sea level at two stations: one in the city centre and another 15 km off-shore, at the "Aqua Alta" platform, which is operated by the Italian Research National Council (CNR). The system provides a good quality forecast, especially on the short time scale. The error of the prediction issued 6 h before the peak of the surge is smaller that 10 cm for 95% of the events (Canestrelli et al., 1999). This skill has not been matched yet by the hydrodynamical models (e.g. Lionello et al., 1998; Bargagli et al., 2002; Cecconi et al., 1997), based directly on fluid dynamics and solving the shallow water equations for current and sea level. This surprising situation is clearly different with respect to other basins, e.g. the North Sea, where the results of the numerical weather prediction models of national meteorological services have been, since the 1980s, routinely used to force hydro-dynamical models for storm surge forecast (Flather and Proctor, 1983 and, in general, Bode and Hardy, 1997). As a result, many countries around the North Sea developed efficient storm surge warning systems, such as the STWS (Storm Tide Warning Service) in the UK (Flather and Proctor, 1983; Flather et al., 1991; Pratt, 1993) and similar systems in The Netherlands (de Vries, 1991; Verboom et al., 1992; Gerritsen et al., 1995) and Denmark (Vested et al., 1992, 1995). The lack of an analogous system for the northern Adriatic has two main scientific explanations. The first one is the errors of the wind field. In the shallow northern Adriatic, the main forcing of the storm surge is the wind stress (Bargagli et al., 2002), which is difficult to predict with sufficient precision by meteorological models, because it presents meso-scale structures produced by the steep mountains on both sides of the Adriatic Sea. The second reason is the sensitivity of the surge prediction to the initial condition. The oscillations of the Adriatic Sea are characterized by a substantially linear dynamics (e.g. Lionello et al., 2005) with a low dissipation, so that seiches triggered by an initial surge persist for several cycles (the attenuation of their amplitude is approximately 10% for each cycle). Consequently, missing a seiche, producing a spurious one, or predicting wrongly its amplitude or phase would spoil the sea level forecast for several days. Dynamical surge models implemented in the Adriatic Sea so far are doomed to propagate the error of the forcing wind field in the surge prediction, without any possibility to compensate for them by a suitable data assimilation procedure. Their prediction, especially on the short (few hours to one day) time range, results worse than that provided by the regressive model, which accounts for the observed sea level. This study aims to improve the forecast provided by a dynamical model, by adopting an adequate variational data assimilation procedure.

The meso-scale pattern, leading to the surge in the northern Adriatic Sea (and the consequent floods of Venice), has already been established in early studies (Robinson et al., 1973). A deep low-pressure system induces a south-eastward pressure gradient along the basin, which, because of the channelling due to the long coastal mountain ridges, produces a strong Sirocco wind. Both wind and, to a lower degree, the inverse barometric effect contribute to surge in the northern part of the basin. The synoptic situations responsible for the surge events have been shown to be originated from a large low pressure system above northern Europe (Trigo and Davies, 2002; Lionello, 2005), which produces a strong orographic cyclogenesis on the southern side of the Alps (Buzzi and Tibaldi, 1978). A further distinction is possible, according to the location of the main minimum with respect to the Alps. The whole cyclone can take a southward path, resulting in an intense cyclone over Italy at the time of the highest surge level, which subsequently attenuates, conserving a well defined identity, during all its evolution. Otherwise, the main pressure minimum can follow a northern path, remaining over central and northern Europe, but originating a secondary minimum south of the Alps, which is subsequently reabsorbed in the deeper system (Lionello, 2005). The occurrence of the surge in the Gulf of Venice is mostly associated to the intensity and frequency of these two synoptic patterns.

This study aims to improve the prediction of a dynamical model by adopting a suitable data assimilation procedure, based on the observed level at the CNR station. The CNR station has been chosen because of both location and data availability. The platform is managed by the Venice CNR and the instrumentation for recording of sea level (with a 10 min. sampling period) is operated by CPSM. Values are transmitted in real time to the CPSM centre itself, where the surge forecast is carried out and a permanent archive of observations and model results is kept. Moreover, the location of the platform 15 km offshore the Venetian coast can be resolved in the grid adopted by the surge model, whose resolution is about 5 km in that area. The measurement is close to Venice and representative of the local sea level, but it is not affected by local features like the details of the coast near the lagoon inlets. Processes important for near-shore values, like coastal set-up, set-down and wave current interactions (which are not included in the model), are not important at that location, where the water depth is about 16 m. Consequently, this measurement is suitable for the assimilation in the model used in this study, that is implemented on the whole Adriatic and, because of its resolution and physics, can provide a forecast of the sea level in the northern Adriatic, eventually close to its coast, but cannot compute levels at the shore, inside the lagoon and its inlets, where a different, dedicated and locally implemented model is a better solution. However, the results of the basin-scale model of this study, could be used to provide the boundary conditions for a local model, describing the currents near the shore or inside the lagoon.

The simulations carried out in this study are based on the hydrostatic padua surface elevation (HYPSE) model, which is a two-dimensional model based on depth averaged currents. Certainly, three-dimensional models provide a more realistic description of bottom friction, which could be very important for properly simulating the interaction between surge and astronomical tide in shallow water. However, tidal amplitude is modest in the Adriatic Sea, and tide and surge can be linearly superimposed without introducing a large error. In other words, the effect of the quadratic bottom friction term and of the advection term in the equation of motion is of minor importance for the simulation of the actual sea level. Moreover, two-dimensional models similar to HYPSE are currently used in state of the art operational systems, when the main purpose is the forecast of the sea level (Flather et al., 1991; Vested et al., 1992; Gerritsen et al., 1995). Finally, the aim of this study is to show that inaccurate forecast in the 1 to 3-day range is mainly due to errors in the initial condition, whose improvement determines large benefits for the forecast, even without accounting for three-dimensional dynamics.

The HYPSE model used in this study is described in Section 2. The data assimilation method is described in Section 3, which includes also the description of the cost function used by the variational method and of the penalty term involved in its expression. The procedure and its effectiveness are described in Section 4. The optimization of the assimilation procedure (involving the choice of weights of the penalty terms and of the length of the analysis period) is described in Section 5. Results are summarized and discussed in Section 6.

2. The model

HYPSE is a standard single layer nonlinear shallow water model, whose equations are derived from the vertical average of the momentum equation, assuming a constant velocity profile. It adopts an orthogonal (eventually curvilinear) C-grid. It uses the leap-frog time integration scheme with the Asselin filter to prevent time splitting. It includes astronomical tide,[3] meteorological forcing (sea level pressure and wind stress), a quadratic bottom stress

$$\vec{\tau}_b = -\rho_w b_f \vec{u}|\vec{u}| \tag{1}$$

and a Smagorinsky horizontal diffusivity with coefficient:

$$A = c_s \Delta x \Delta y \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 + \frac{1}{2}\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right)^2 \right], \tag{2}$$

where u, v are the depth-averaged current velocity components, Δx and Δy are the grid steps, b_f and c_s are two coefficients, whose value has been chosen to optimize the model results, ρ_w is the water density.

The model implementation used in this study has been selected in order to optimize the model results and achieve high resolution in the northern part of the basin, while limiting the computer resources needed for the simulations. The dependence of the model accuracy on the grid characteristics and resolution has been investigated and three different grids have been tested (Lionello et al., 2005). The best results are obtained using a rectangular mesh grid of variable size, which has the highest resolution in the northern part of the basin at 14°E, 44°N, where the minimum step is 0.03 degrees. Starting from that point, the grid step increases with a logarithmic increment (which uses a 1.01 factor) in both latitude and longitude. Practically, its resolution varies in the range from 3.3 to 7 km. The whole Adriatic Sea is covered with $N_x = 133$ and $N_y = 146$ points in longitude and latitude, respectively. In Fig. 1, the left panel shows the area covered by the grid, the right panel shows a detail with the grid points in the northern part of the Adriatic Sea.

The optimal values of b_f, bottom friction coefficient, and c_s, Smagorinsky diffusivity coefficient, have been identified by comparing the astronomical tide, simulated by the model, with the observed amplitude at 9 stations along the coast of the Adriatic Sea (Fig. 2). Fig. 3 shows the root mean squared (RMS) error of the model results considering all nine stations as function of b_f and c_s. The minimum RMS error of 1.4 cm is obtained at $b_f = 5 \times 10^{-3}$ and $c_s = 0.4$. Fig. 4 shows the RMS error values at the single stations.

In the surge simulations HYPSE is forced by the sea level pressure (SLP) and surface wind stress $\vec{\tau}_W$ fields. The stress is parameterized as

$$\vec{\tau}_W = \rho_a C_D U_{10} \vec{U}_{10}, \quad C_D = A + BU_{10}, \tag{3}$$

where $\vec{U}_{10}$ is the wind speed at the 10 meter level, and $A = 8 \times 10^{-4}$, $B = 8.4 \times 10^{-5}$ s/m if $U_{10} > 4.8$ m/s, and $A = 1.2 \times 10^{-3}$, $B = 0$ otherwise. These values of A and B are meant to account for increased drag, due to the effect of the wind waves in the strong wind conditions, which produce the storm surge in the northern Adriatic Sea (Lionello et al., 1998). The 6-hourly SLP and $\vec{U}_{10}$ fields were provided by the European Centre for Medium range Weather Forecasts (ECMWF) analysis at T213 spectral resolution. Both the associated space resolution (about 50 km) and time resolution are often not adequate for capturing the sharp features responsible for the surge events, therefore the need

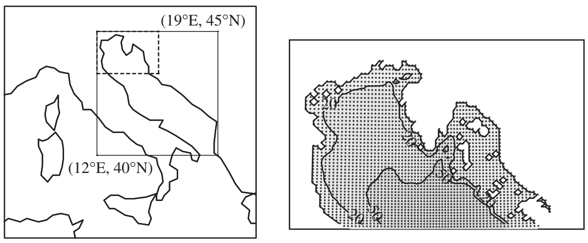

Fig. 1. Left panel: the box shows the area covered by the model grid (corners are denoted with the corresponding coordinates). Right panel: grid detail in the northern part of the basin (corresponding to the dashed squared box in the left panel). The dots represent the grid points.

[3]Seven tidal constituents were used in the simulations considered in this study.

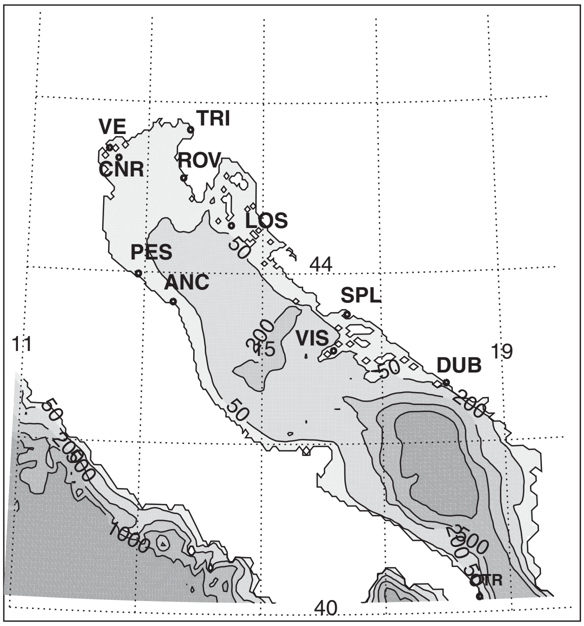

Fig. 2. Bathymetry of the Adriatic Sea (values in metres) and location of the stations used for the validation against tide: TRI Trieste, CNR Venice (CNR platform), PES Pesaro, ANC Ancona, SPL Split, DUB Dubrovnik, LOS Losinj, ROV Rovinj, VIS Vis. The data at Otranto OTR were not used for the validation but for the boundary condition at the southern open border of the basin. The location of the Venice city centre is denoted as VE. The longitude and latitude step is drawn with a 2 degs step, starting from 11W and 40N.

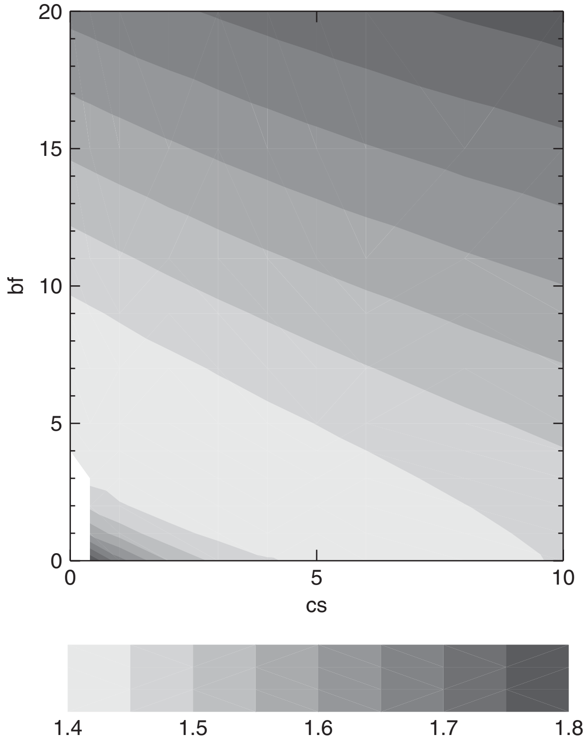

Fig. 3. Accuracy of the model for tidal simulation. The plot shows the RMS error (cm) on the basis of 9 stations along the coast of the Adriatic Sea. The RMS error is computed as function of the Smagorinsky constant (x-axis, $c_s \times 10$) and of the bottom friction coefficient (y-axis, $b_f \times 10^4$). Contour levels from 1.40 to 1.80 cm, contour interval 0.05 cm. The white area denotes parameter values where the model is unstable.

for data assimilation. The boundary condition at the southern open boundary of the Adriatic was based on the sea level observed at the Otranto station (Fig. 2). Consequently, the simulations were carried out in hindcast mode. Of course, analyzed meteorological forcing and observed sea level boundary conditions would not be available for an operational forecast, but the organization adopted for the simulations in this study allows to isolate the error associated to the initial condition. In the operational practice, the error of the meteorological forecast and the lack of information on the sea level value at the open end of the Adriatic would also contribute to the error.

3. The data assimilation method

The assimilation procedure is based on the adjoint model (Lewis and Derber, 1985; Talagrand and Courtier, 1987; Thacker and Long, 1988; Thacker, 1988), in order to compute the gradient of the cost function with respect to the initial condition. In general, the cost function J is given as

$$J = J_O + J_P = \sum_{i=1}^{N_t} (J^i_O + J^i_P). \tag{4}$$

The terms J^i_O and J^i_P represent the contribution to the cost function at time level t^i, where

$$J^i_O = \sum_{p=1}^{N_{\mathrm{Obs}}} (\eta^i_{k_p l_p} - \eta^i_p)^2 \tag{5}$$

if observations are available at time t^i,

$$J^i_O = 0 \quad \text{otherwise, and}$$

$$J^i_P = w_1 P^i_1 + w_2 P^i_2 + \cdots + w_M P^i_M, \tag{6}$$

where N_t represents the number of integration steps Δt, $\eta^i_{k_p l_p}$ represents the sea surface elevation at the grid point closest to station (no space interpolation has been carried out), η^i_p the observed value, i the

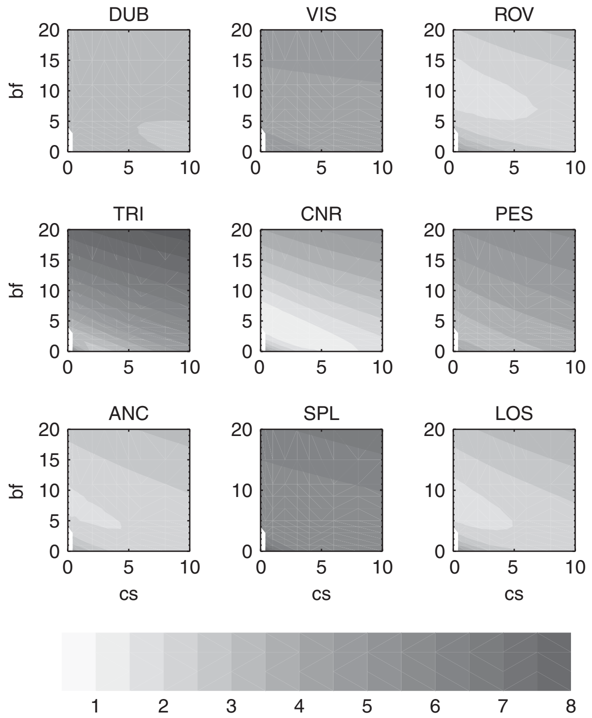

Fig. 4. Accuracy of the model for tidal simulation. The plots show the RMS error (cm) for the 9 stations in the map of Fig. 2. The RMS error is computed as function of the Smagorinsky constant (x-axis reports $c_s \times 10$) and of the bottom friction coefficient (y-axis reports $b_f \times 10^4$). Contour levels from 0.5 to 8.0, contour interval 0.5 in all panels. The white area denotes parameter values where the model is unstable.

time index, $t^i = t^0 + i\Delta t$. The symbols P_m^i represents the M penalty term at time t^i, whose utility will be discussed later, which are used to ensure stability of the minimization and physically meaningful solution (Thacker, 1988), and w_m, $m = 1, \ldots, M$ are dimensionless weights, controlling the strength of their influence. In this study only one penalty term ($M = 1$) and one observation ($N_{\mathrm{Obs}} = 1$) at the CNR platform η^i_{CNR} have been inserted in the expression used for the operational implementation.

In this formalism, the state of the system at time t_i, is represented by a $3N$-dimensional vector $\vec{X}^i$, with N number of grid points, which consists of the two components of the current velocity and the elevation, so that $\eta^i_{k_{\mathrm{CNR}} l_{\mathrm{CNR}}}$, model elevation at the CNR station, is a single element of $\vec{X}^i$.

The adjoint model has been used to improve bathymetry and open boundary conditions in surge and tidal models (Lardner et al., 1993; ten Brummelhuis et al., 1993). The Kalman filter has been used to assimilate sea surface observations and correct the state of the system while integrating forward in time, thereby establishing initial conditions which are optimal in a statistical sense (Gerritsen et al., 1995; Canizares et al., 1998). The procedure described in this study finds the optimal initial condition by identifying the model simulation which best reproduces the observations during the analysis period prior to the forecast. This solution is particularly suitable for the Adriatic Sea, because of the large amplitude and small attenuation of the seiches characterizing its dynamics, which has no analogy in other basins, like the North Sea.

The adjoint method considers the linear version of the model, which has been obtained by neglecting the nonlinear advection term, replacing the quadratic bottom friction with

$$\vec{\tau}^l_b = -\rho_w b^l_f \vec{u}, \tag{7}$$

with $b^l_f = 7 \times 10^{-4}\,\mathrm{m/s}$, and keeping a constant horizontal diffusivity coefficient, $A^l = 500\,\mathrm{m^2/s}$.

The linear model forward integration in time can be represented by a matrix M, with elements M^i_{km}, and a sequence of multiplications, each advancing the solution in time for a model time step Δt, at time t^i

$$X^{i+1}_k = X^i_k + \Delta t M^i_{km} X^i_m, \tag{8}$$

while the computation of the gradient of the cost function J with respect to the variable X^{i-1}_k at time level $i-1$, $D^{i-1}_k = \partial J / \partial X^{i-1}_k$, is given by the sequence

$$D^{i-1}_k = D^i_k + \Delta t A^i_{km} D^i_m + \frac{\partial J_O}{\partial X^{i-1}_k} + \frac{\partial J_P}{\partial X^{i-1}_k}, \tag{9}$$

representing an integration backward in time of the adjoint A of M, with $A^i_{km} = {M^i_{mk}}^*$. Eq. (9), at difference with the linear model equation (8), contains a forcing represented by the gradient of the cost function contribution, at time level i, with respect to the model variables at the same time level. The application of this chain rule, from the final step of the simulation to the first one, gives the gradient of the cost function J, with respect to the initial condition $\vec{X}^0$.

The data assimilation procedure iterates single steps while moving towards the optimal initial condition. During each iteration the cost function, by forward integration of the linear model, and its gradient, by backward integration of the adjoint model, are computed and the conjugate gradient method is used for identifying the optimal direction along which to search for a minimum in the $3N$-dimensional space. A parabolic profile for the surface describing the cost function in the $3N$-dimensional space is assumed and the so-called "golden ratio rule" is used for estimating the location of the minimum (Press et al., 1992) along this conjugate direction. The following iteration step starts from this point and the procedure continues until when a local minimum is found.

The procedure has been tested in a synthetic case. A model simulation has been used for generating a seiche and the hourly values of the sea surface elevation at the grid point corresponding to the CNR platform have been used as a synthetic set of data. In practice, the model was forced imposing an oscillation at the southern border of the Adriatic Sea, across the Otranto Strait. After 30 days, when the oscillation had reached a steady state in which its amplitude did not change anymore, the forcing at Otranto was stopped, so that the seiche began decreasing its amplitude. Hourly data during the first 3 days, after switching off the forcing (and during which the amplitude of the seiche was decreasing), have been used as artificial observations to be assimilated by the procedure. No penalty term was included in the cost function for this test. The aim of the assimilation procedure is to reconstruct the decay of the seiche, beginning from a state of rest ($\vec{X}^0 = 0$) as first guess. Fig. 5 shows the adjustment of the solution to the model synthesized observations (top panel) and the reduction of the cost function with the number of iterations (bottom panel). A maximum of 25 iterations was allowed both in this synthetic case and in the realistic experiment described in the next session. In this simple case, in which no further source of error (as a wrong meteorological forcing or model errors) is involved, the assimilation procedure improves the initial condition and a very good reproduction of the seiche is obtained.

4. The data assimilation and surge prediction procedure

The operational prediction procedure has a relatively standard structure, according to the sequence shown in Fig. 6. Every day one simulation, split into an analysis and a forecast period, is carried out. During the analysis, the model is forced by analyzed surface wind and SLP fields and the procedure assimilates the observations, in order to provide the optimal initial condition for the forecast. The first iteration begins from a state of rest with null elevation and current. The initial condition of the following iteration are obviously determined by the progress of the data assimilation procedure. The adopted duration of the analysis is 3 days, though also the effect of a different length has been tested (see following discussion). During the forecast period, whose length has been limited to 3 days in this study, but it could be easily extended, the model is forced by surface wind and SLP fields without any further correction.[4] The forecast period corresponds to that of the actual prediction, produced by the model, and it would be normally delivered about 8 h after its initial time, such delay being produced by the time required for the transfer of the forecast forcing fields to the local computing resources and for the completion of the data assimilation procedure and the forecast. The computational cost is low and poses no problem for operational application. The whole assimilation, consisting of a maximum of 25 iterations, is completed in about 10 h and the prediction takes less than 5 min using a single processor Alpha EV67@667Mhz.

This study considers a series of experiments, each of them consisting of a series of simulations split in an analysis and a forecast period, whose aim is to provide a forecast of sea level. In the experiments carried out in this study, the operational procedure has been applied to a sequence of daily simulations for the period 5th to 24th November 1996, during which a sequence of large surge events took place.

A parallel set of CTR_n (ConTRol) experiments have been carried out, initializing the forecast with a n-day long spin-up period from a state of rest, with n varying in the range from 0 to 8, and not using the data assimilation procedure.[5] During the spin-up period, the model was forced by the ECMWF analysis and assumed a state of rest as initial condition. In the organization of the CTR_n experiments, the spin-up period corresponds to the analysis period of the assimilation experiments. Also a CTR experiment, consisting of a single continuous simulation, covering the whole period has been carried out. The comparison between assimilation and CTR_n experiments is used to evaluate the benefits of the assimilation procedure. Table 1 shows the RMS error, bias and standard

[4]In the experiments carried out in this study, analyzed SLP and surface wind fields are used. In the actual operational prediction, forecast fields would be used.

[5]The label 0 denotes the experiment without spin-up, meaning a forecast with a state of rest as initial condition.

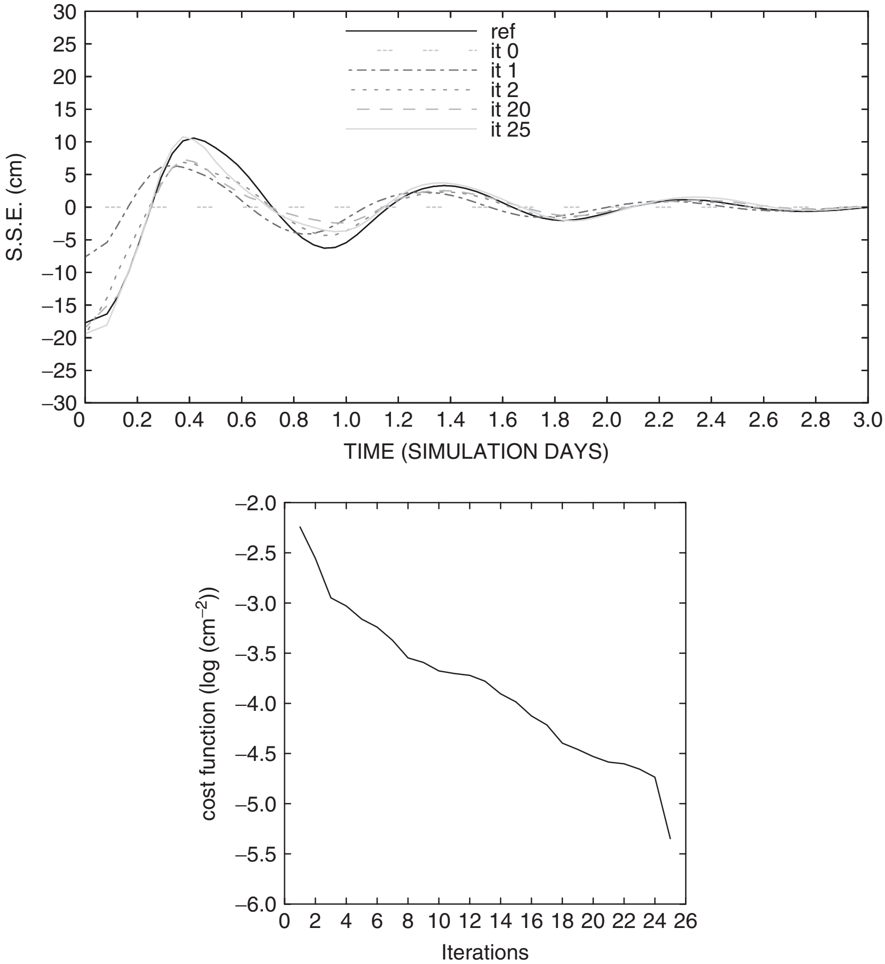

Fig. 5. Top panel shows the adjustment of the model solution at the CNR platform to the model synthesised time series (denoted with "ref") used for the assimilation test for iteration 0 (state of rest), 1, 2, 20 and 25 when the procedure was stopped. The bottom panel shows the reduction of the cost function (value in cm^2 on y-axis) with the number of iterations (x-axis).

deviation, considering the 0–24 h time window of the forecast produced by the CTR_n experiments. Results show that the quality of the forecast improves with the duration of the spin-up, but no further benefit is obtained if it is extended beyond 4 days. Fig. 7 has a similar information, but it resolves the behaviour of the RMS error in time during the period considered. The first part of the simulated period (from 5th to 13th) was characterized by the absence of relevant events and no significant error, which grows during the second part, during which relatively high surge events took place. Fig. 7 shows that the benefit of extending the spin-up period are rather constant in time, larger for the 0–24 h forecast period, and confirms that a duration longer than 4 days has no further effect. The information from Fig. 7 and Table 1 supports an optimal 5-day duration for the spin-up period. Therefore, the CTR_5 experiment provides a reference for evaluating the benefits of the assimilation.

The sequence of events from the 15th to the 24th can be identified in Fig. 8, which shows the observations, the CTR_5 simulation, and three sequences of simulations, extracted from the assimilation experiment and consisting of analysis and forecast with initial date 18th, 19th, 20th November 1996. Only the storm surge contribution to the sea level is considered. The highest storm surge took place on the 18th, followed by other peaks and anticipated by a sequence of smaller storm surges. The results of the model without data assimilation (the CTR_5 experiment) are unsatisfactory, because of recurrent and severe underestimation of the peak levels. The time series shows the improved

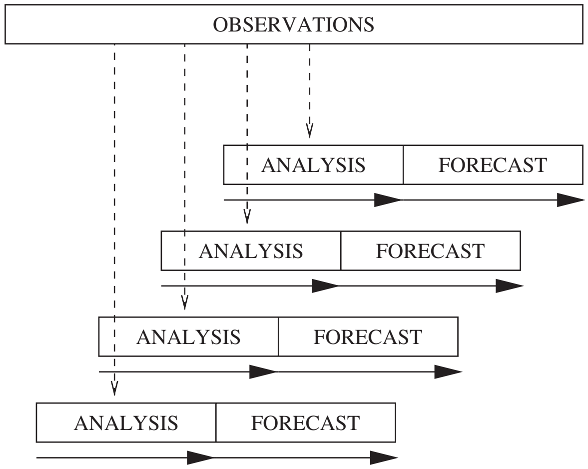

Fig. 6. The scheme shows the sequence of operations for the whole experiment. Each day an analysis and a forecast are carried out, providing a prediction in the range up to 3 days.

Table 1
RMS error of the surge simulations for the whole simulate period. Rows 1–4 refer to the "O + P" experiment and consider the last day of the analysis, first, second and third day of the forecast, respectively. Rows 5–10 refer to the first day of the forecast in the "CTR_n" experiments, where n varies from 0 to 5. The time series analyzed cover the period between 07 UTC of 13th November and 06 UTC of 23rd November. Values in centimetres. A further extension of the spin-up period does not change appreciably the results

Time series	CNR platform	Trieste
Analysis	5.7	14.5
1day forec.	12.7	19.8
2day forec.	14.2	19.9
3day forec.	17.4	21.8
CTR_0	28.9	32.3
CTR_1	25.5	28.4
CTR_2	24.1	26.7
CTR_3	23.3	25.8
CTR_4	23.1	25.7
CTR_5	23.1	25.7

agreement of model and observations obtained by the data assimilation during the analysis period (when the model results overlap almost exactly the observations) and the progressive deterioration of the results with the time range of the forecast, which, however, improves with respect to the CTR_5, also in the 3-day time range. The assimilation experiment adopted a 3-day duration for the analysis period and the cost function obtained adding to the model error a penalty term P_{ns},

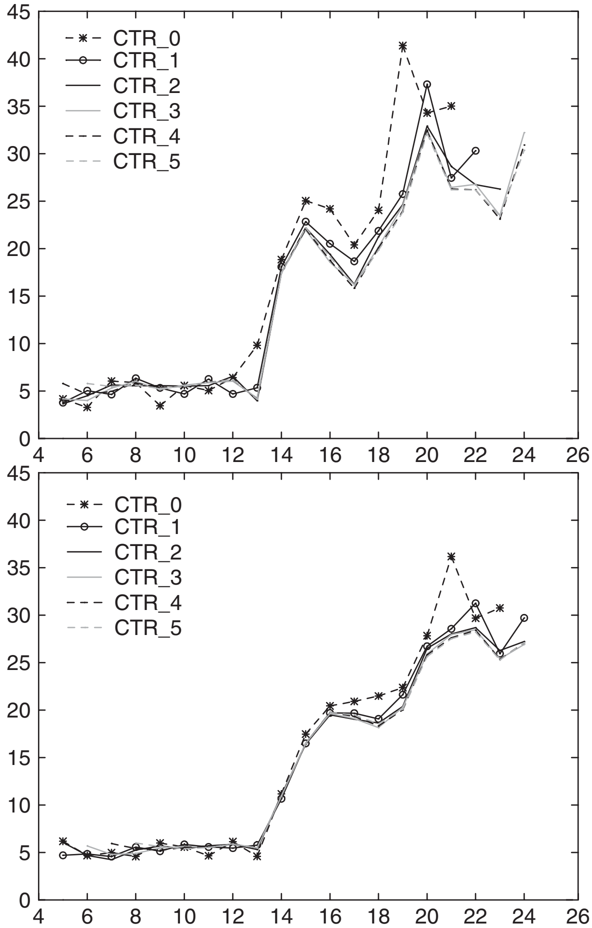

Fig. 7. Dependence of the RMS error (y-axis, cm) of the forecast on the length of the spin-up period of the CTR_n experiments. Top panel: 0–24 h forecast. Bottom panel: 0–72 h forecast. The x-axis shows the initial date of the forecast. The time window considered for the RMS computation ends at 06 UTC of the day on x-axis.

whose characteristics and effect are described in Section 5, hereafter.

The benefits of the assimilation are confirmed by Table 1. The RMS error in sea level is small during the last day of the analysis and it increases with the range of the forecast. However, the RMS error of the 48–72-h forecast with the assimilation is still lower than that of the 0–24 h of all CTR experiments. The benefits of the assimilation are confirmed in Fig. 9, which shows the time behaviour of the RMS error, with respect to the CTR_5 experiment, for different values of the forecast range. The benefits decrease in time, but persist also in the 3rd day of the forecast. Fig. 10 shows the effect of the data assimilation, with respect to the CTR_5 experiment. The four time series of the data

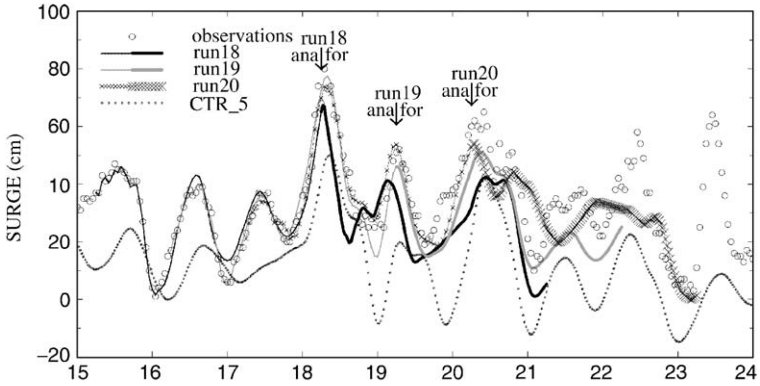

Fig. 8. Observations (circles) and results of the data-assimilation experiment (values in cm). Dots denote the CTR_5 experiment. The other curves show three sequences of analysis-forecast from the data assimilation experiment with initial date 18th (black line), 19th (gray line), 20th (crosses) November 1996. For each sequence an arrow indicates the separation between analysis and forecast period (which is denoted with a thicker line). Only the storm surge contribution to the sea level is shown. The time series show the excellent agreement of model and observations during the analysis period and the progressive deterioration of the results with the time range of the forecast.

assimilation experiment have been obtained joining 1-day long parts extracted from the sequence of simulations, so that the time series labelled "0–24 h", "24–48 h", and "48–72 h" comprise the first, second and third day of the forecast, respectively. The time series labelled "analysis" is composed by the last day of the analysis period of each simulation. The different simulations used the same analyzed wind field and differences in the time series derive only from the initial condition. Fig. 10 confirms that the analysis follows very closely the observations. Though the error increases with the forecast range, a clear improvement is present also in the 48–72 h forecast (only occasionally the RMS error can be lower in a 3-day than in a 2-day forecast). The results show that the data assimilation procedure is capable of improving the forecast in the range up to 3 days (and eventually further in time).

Though only the hourly values at the CNR platform are assimilated, the improvements on the model results are not restricted to the Venetian shore. Table 1 shows also the model statistics in Trieste, where no observations were assimilated. With respect to the CTR_5 simulation the RMS error is reduced more than 50% during the analysis, and the improvement persists until the end of the 3-day forecast, decreasing with the forecast range as it does at the CNR platform.

Finally, the influence of the length of the analysis period has been explored. Four assimilation experiments, considering the sequence of three predictions, starting on the 18th, 19th, 20th November

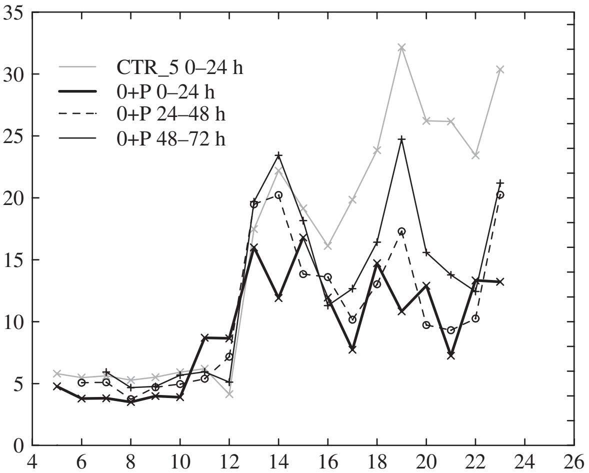

Fig. 9. Time behaviour of the daily RMS error for the CTR_5 experiment (0–24 h forecast period, gray curve) and for the data assimilation experiment considering the 0–24 h, 24–48 h, and 48–72 h forecast range.

1996, have been carried out with varying the length of the analysis from 1 to 4 days. Fig. 11 shows that an 1-day duration is fully inadequate, 2-day is insufficient and the best forecast is obtained with a 3-day duration. A further extension to 4 days actually spoils the forecast. It is likely that a short window does not include sufficient data for adapting the initial condition to the observed sea level oscillation, while a too long window moves too far backward in time the initial condition with respect to the initial time of the forecast.

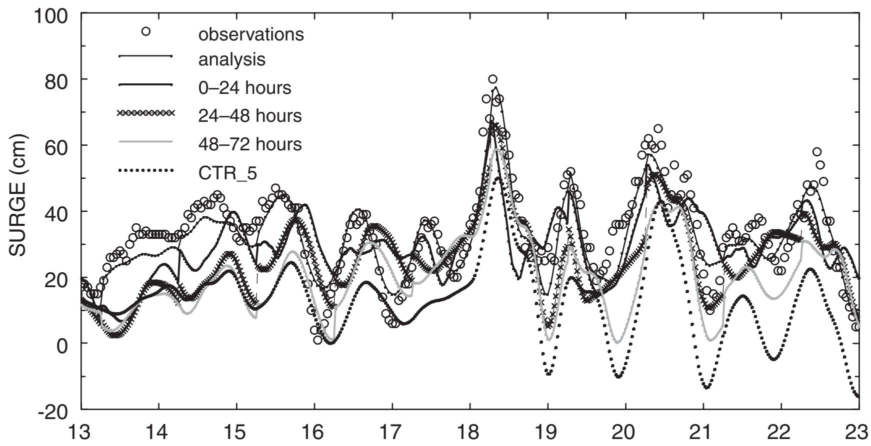

Fig. 10. Central part of the period during which the operational data assimilation has been carried out. Only the storm surge is shown. Figure shows the observations (circles), the time series labelled "analysis" (composed by the last days of the analysis period of each simulation), the time series labelled "0–24 h", "24–48 h", and "48–72 h" (composed by the first, second and third day of the forecast in the data assimilation experiment, respectively) and the CTR_5 simulation.

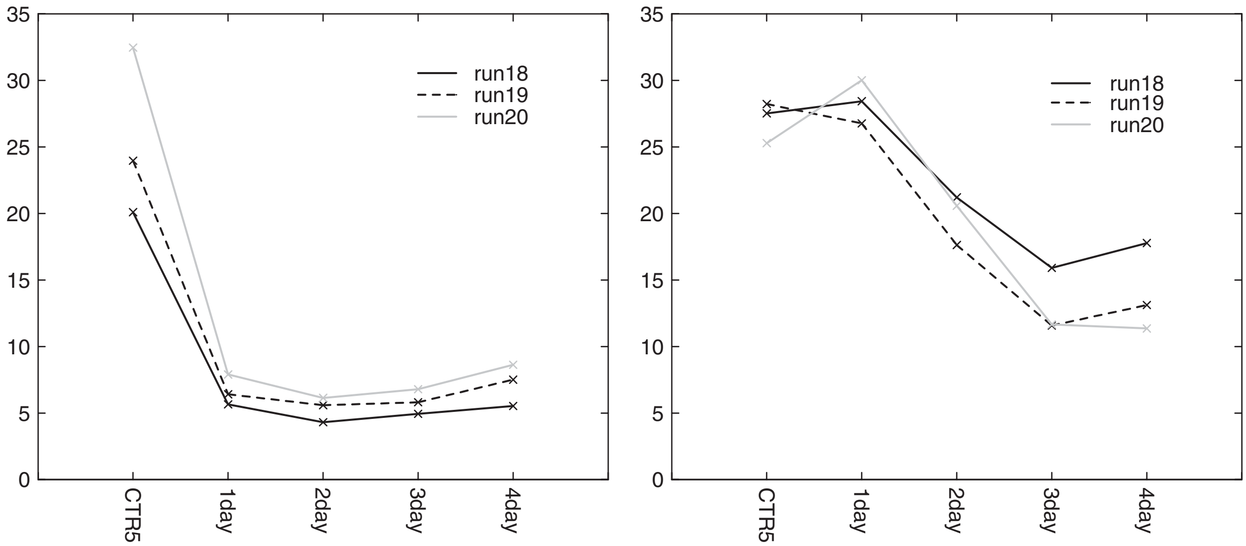

Fig. 11. Statistics of "O + P" experiments as function of the length of the analysis period (from 1 to 4-day duration, *x*-axis) with respect to that of the CTR_5 experiment. The panels show the RMS error (centimetres) of the last day of the analysis (left panel) and of the subsequent 3-day long forecast (right panel).

5. The importance of the penalty term

There are some problems associated with an optimization of the initial condition, based only on the single observation at the CNR platform. At this location, first (22-h) seiche, second (11-h) seiche and surge, which are the main components of the sea level oscillations, have a large amplitude, so that observations show clearly their presence in the basin (Lionello et al., 2005). However, this single observation is unable to account for the behaviour of the sea surface in the rest of the basin, where spurious oscillations and short scale waves are allowed to grow, and could spoil the forecast when they reach the northern part of the Adriatic. Moreover, their amplitude grows, in general, at the coast, particularly in the narrow and shallow gulfs along the eastern coast, and they could introduce numerical instabilities.

The insertion of the penalty terms J_P in Eq. (4) is needed to stabilize the data-assimilation procedure. Its purpose is to avoid that the procedure considers initial conditions with strong unbalance between current and elevation, small and steep elevation features, large local divergence of the current field. The presence of these unrealistic features in the initial condition would not correspond to an optimal initial state (as the cost function is concerned), and would be discarded during the assimilation, but their occurrence during the search for the minimum can determine numerical problems, instabilities and the premature interruption of the assimilation procedure. The inspection of the two-dimensional sea surface elevation field shows that short scale "noise" responsible for this problem is originated in the southern deep part of the Adriatic and it produces large instabilities, when it reaches the coast. It is likely that the availability of measurements to constrain the simulation in that part of the domain could eliminate the problem. However, such information is not presently available for this operational implementation and to use a penalty term is necessary.

Different forms of the penalty term have been considered and tested in order to identify a reliable expression for the cost function. The test considers the central part of the analyzed period, when the highest surge peaks took place, and it consists of the sequence of three predictions, starting on the 18th, 19th, 20th November 1996. In practice, all terms penalising high frequency and steep waves are somehow successful and their results are comparable. The term P_{ns} has been chosen for the operational implementation as it resulted particularly reliable and robust for reducing noise over the whole model domain. It is proportional to the squared Laplacian, averaged on the whole duration of the analysis

$$J_P = 10P_{ns},$$

$$P_{ns} = \frac{1}{N_t N_S}\sum_{i=1}^{N_t}\sum_{k,l}^{N_S} \times [4\eta^i_{k,l} - \eta^i_{k+1,l} - \eta^i_{k-1,l} - \eta^i_{k,l+1} - \eta^i_{k,l-1}]^2, \quad (10)$$

where sum with indexes k, l is carried out on all N_S sea points of the grid.

In the following discussion and figures the label "O" denotes the experiment where only the term J_O was used and the label "O + P" that where also the penalty P_{ns} was introduced in the cost function.

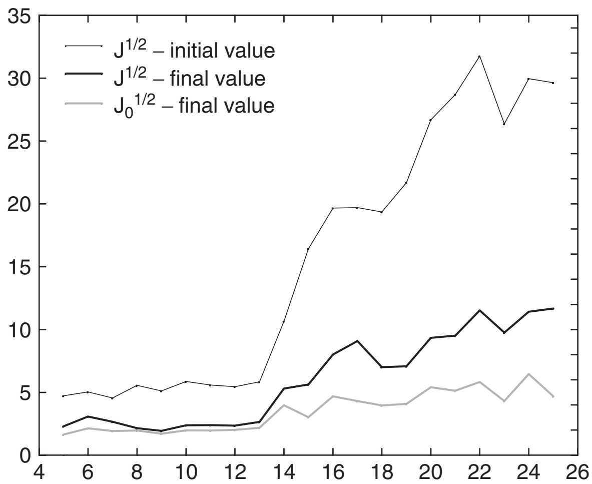

Fig. 12. Efficiency of the procedure in reducing the RMS error during the 3-day analysis in the "O + P" experiment. $J^{1/2}$ is the squared root of the whole cost function (the thin and thick lines show the value before and after the data assimilation procedure), and $J_O^{1/2}$ the RMS error. Initially, the penalty term J_P has a very small value, so J is practically identical to J_O, whose initial value is not shown in this figure.

The three curves in Fig. 12 show the squared root of the initial and final values of J, that is at the beginning of the iteration procedure and after the completion of the assimilation, and the final RMS error. The evolution of these quantities during the whole simulated period is considered. The RMS error (not shown) before the assimilation is practically identical to $(J)^{1/2}$, because J_P is negligible during the first iteration, and increases during the assimilation procedure and can be comparable to that of J_O at the end of the assimilation. This happens during the second part of the simulated period (after the 14th), when the assimilation increases the energy of the initial condition for improving the reproduction of a sequence of high surge events.

However, the introduction of P_{ns} does not appreciably affect the reduction of the RMS error. Fig. 13 shows the results of the data assimilation for the whole period and compares the result of experiments "O" and "O + P" and "CTR_5", by showing the daily value of the residual RMS error, during the last day of the analysis (left panel) and during the whole 3-day long forecast period (right panel). The number of iterations, carried out during the "O" experiment, is also shown in the figure. It shows that the two assimilation experiments produce equivalent results when "O" is capable of completing an acceptable number of iterations, but

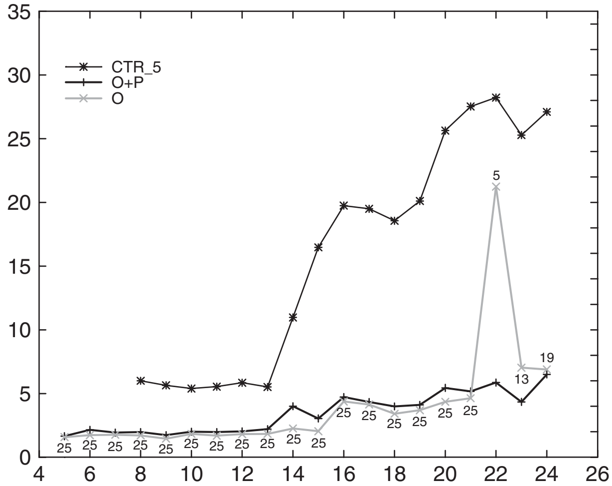

Fig. 13. Reduction of the RMS error during the last day of the analysis period (left panel) and during the 3-day long forecast period (right panel) obtained by the data assimilation (experiments "O" and "O + P") with respect of the CTR_5 experiment.

Table 2
RMS error during the last day of the analysis (second column), its reduction with respect to the CTR_5 experiment (third column), RMS error during the subsequent 3-day long forecast (fourth column), and its reduction with respect to the CTR_5 experiment (fifth column) for the "O", "O + P", and the CTR_5 experiments. Values (in centimetres) represent overall results for simulations with forecast starting on 18th, 19th, 20th November 1996

Experiment	Analysis	% Analysis	Forecast	% Forecast
O	3.8	17.2	11.7	43.3
O + P	4.6	20.8	13.2	48.9
CTR_5	22.1	100.0	27.0	100.0

"O" fails to continue the iterations when critical situations, which can develop because of an unrealistic initial condition met during the search for the minimum, determine a numerical instability in the forward integration. The "O + P" experiment always executes 25 iterations and completes the minimization with good results.

The residual RMS error, after the assimilation is completed, is shown in Table 2. The introduction of the penalty term does not appreciably influence the reduction of the RMS error. With respect to the CTR_5 experiment, a very large reduction is obtained by both "O" and "P + O" during the analysis, during which the residual error is about 20% of that in the CTR_5 experiment, and an important reduction persists for the whole 3-day long forecast, during which the residual error is about 50% of that in the CTR_5 experiment. Actually, the penalty term introduces extra constraints and, consequently, tends to allow a marginally larger residual error. On this respect it slightly reduces the efficiency of the assimilation procedure. In fact, the advantage of the penalty term is not to provide a more efficient reduction of the RMS error, but to increase the stability of the procedure.

6. Conclusions

This study shows that a variational procedure is capable of efficiently improving the results of a storm surge prediction model in the Adriatic Sea, by optimizing the initial condition of the forecast. The procedure is based on the minimization of a cost function, which adopts the adjoint model for the computation of its gradient with respect to the initial condition. A penalty term is used, to avoid the presence of spurious steep small scale features and increase stability of the procedure.

The procedure compensates for a limitation of the dynamical models implemented so far in the Adriatic, which were unable to account for observed values and correct the model results. The data assimilation greatly improves the prediction, whose quality is useful for operational applications such as the management of movable barriers planned to protect Venice and the lagoon and for providing an early warning for the city centre management, in order to avoid loss of goods and troubles for the population. Improvements of the forecast, due to the data assimilation, are consistently large in the range up to 3 days. A future study, covering a longer period and carried out in a fully operational framework at CPSM, is planned, in order to assess the benefits in a truly operational framework.

Further improvements to the operational prediction procedure are certainly possible and recommended. They are related to the insertion of other observations in the central and southern part of the basin (e.g. the data that could be available in real time at Ancona), and to the extrapolation forward in time of the level at the southern boundary of the basin, across the Otranto strait. While in all simulations of this study, which have been carried out in hindcast mode, it was possible to impose the observed level, instead the operational forecast would benefit from a procedure capable of providing a boundary condition and accounting both for the changing mean level at the southern boundary and the presence of periodic oscillations. Finally, and presumably most importantly, to devise a scheme capable of consistently modifying wind fields and model initial condition would greatly improve the prediction. The problem is the correction of the wind during, on one hand, the analysis, on the other hand, the forecast.

Though improvements of the meteorological forcing, which is commonly blamed for the poor performance of surge models, are certainly important, this study shows that also a relatively cheap method, aiming at an improved initial condition, is capable of producing large benefits to the forecast, reducing its error to 50% of the original value in the 1 to 3-day forecast range. In fact, because of the low dissipation and the large seiches characteristic of the Adriatic Sea, the accumulation of errors, derived from previous simulations in the initial condition, is the largest source of error for the surge prediction.

Acknowledgements

The hourly sea level observations were kindly provided by "Centro Previsioni e Segnalazioni Maree" of the Venice City Council. The study was partially funded under contract between the University of Padua and the Venice City Council and it was made possible by the support of the University of Lecce. The authors are indebted to Prof. A. Tomasin for his help with analysis of sea level data and surge level computation.

References

Bargagli, A., Carillo, A., Mariotti, A., Pisacane, G., Ruti, P.M., Struglia, M.V., 2002. An integrated forecast system over the Mediterranean basin: extreme surge prediction in the northern Adriatic Sea. Monthly Weather Review 130, 1317–1332.

Bode, L., Hardy, T., 1997. Progress and recent developments in storm surge modeling. Journal of Hydraulic Engineering 123, 315–331.

Bondesan, M., Castiglioni, G.B., Elmi, C., Gabbianelli, G., Marocco, R., Pirazzoli, P.A., Tomasin, A., 1995. Coastal areas at risk from storm surges and sea-level rise in Northeastern Italy. Journal of Coastal Research 11, 1354–1379.

Buzzi, A., Tibaldi, S., 1978. Cyclogenesis on the lee of Alps: a case study. Quarterly Journal of the Royal Meteorological Society 104, 171–287.

Canestrelli, P., 1999. Il sistema statistico del Comune di Venezia per la previsione del livello della marea in città. Risultati teorici e in fase di operatività. Comune di Venezia—CPSM, unpublished.

Canestrelli, P., Cecconi, G., Di Donato, M., 1999. Valutazione dell'accuratezza dei livelli di acqua alta a Venezia ottenuti con il prototipo del sistema di previsione del periodo 1997–1998. In: Quaderni Trimestrali, Consorzio Venezia Nuova, Venezia, vol. 1, pp. 5–12.

Canestrelli, P., Mandich, M., Pirazzoli, P.A., Tomasin, A. 2001. Wind, depression and seiches: tidal perturbations in Venice (1951–2000). Comune di Venezia—CPSM, unpublished.

Canizares, R., Heemink, A.W., Vested, H.J., 1998. Application of advanced data assimilation methods for initialisation of storm surge models. Journal of Hydraulic Research, ISSN 002-1686 26 (4), 655–674.

Carminati, E., Doglioni, C., Scroccas, D., 2005. Long term natural subsidence of Venice: evaluation of its causes and magnitude. In: Fletcher, C., Spencer, T. (Eds.), Flooding and Environmental Challenges for Venice and its Lagoon: State of Knowledge 2003 Meeting (Cambridge, September 2003). Cambridge University Press, Cambridge, UK, pp. 21–28.

Cecconi, G., 1999. Real time storm and watershed inflow forecasting in the Venice lagoon. In: Integrated system for real time flood forecasting and warning, RIBAMOD Workshop, (Monselice, September 25–26th 1997), Luxembourg, 1999, pp. 147–163.

Cecconi, G., Canestrelli, P., Di Donato, M., Cecchinato, M., 1997. I nuovi sistemi per la previsione dell'acqua alta. In: Quaderni Trimestrali, Consorzio Venezia Nuova, Venezia, 3/4, pp. 23–39.

de Vries, J.W., 1991. The implementation of the WAQUA/CSM-16 model for real time storm surge forecasting. Technical Report TR-131, KNMI.

Flather, R.A., Proctor, R., 1983. Prediction of North Sea storm surges using numerical models: recent developments in the U.K. in North Sea dynamics. In: Sundermann, J., Lenz, W. (Eds.), North Sea Dynamics. Springer, Berlin, pp. 299–317.

Flather, R.A., Proctor, R., Wolf, J., 1991. Oceanographic forecast models. In: Farmer, D.G., Rycroft, M.J. (Eds.), Computer modelling in the environmental sciences. Clarendon Press, Oxford, pp. 15–30.

Frassetto, R., 1976. Altimetria del centro storico di Venezia. Bollettino di Geofisica Teorica ed Applicata, vol. XVIII, n. 71.

Gerritsen, H., de Vries, H., Philippart, M., 1995. The Dutch Continental Shelf Model. In: Lynch, D., Davies, A. (Eds.), Quantitative Skill Assessment for Coastal Ocean Models, Coastal and Estuarine Studies, vol. 47. American Geophysical Union, Washington, USA, pp. 425–467.

IPCC 3rd Assessment Report Climate Change, 2001. The Scientific Basis Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). In: Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., van der Linden, P.J., Xiaosu, D. (Eds.), Cambridge University Press, Cambridge, UK, 944p.

Lardner, R.W., Al-Rabeh, A.H., Gunay, N., 1993. Optimal estimation of parameters for a two-dimensional hydrodynamical model of the Arabian Gulf. Journal of Geophysical Research 98 (C10), 18229–18242.

Lewis, J., Derber, J., 1985. The use of adjoint equations to solve a variational adjustment problem. Tellus 37A, 309–322.

Lionello, P., 2005. Extreme surges in the Gulf of Venice: present and future climate. In: Fletcher, C., Spencer, T. (Eds.), Flooding and Environmental Challenges for Venice and its Lagoon: State of Knowledge 2003 Meeting (Cambridge, September 2003). Cambridge University Press, Cambridge, UK, pp. 59–70.

Lionello, P., Zampato, L., Malguzzi, P., Tomasin, A., Bergamasco, A., 1998. On the correct surface stress for the prediction of the wind wave field and the storm surge in the northern Adriatic Sea. Il Nuovo Cimento C 21, 515–532.

Lionello, P., Mufato, R., Tomasin, A., 2005. Effect of sea level rise on the dynamics of free and forced oscillations of the Adriatic Sea. Climate Research 29, 23–39.

Pratt, I., 1993. Operation of the Storm Tide Warning Service. In: Proceedings of the 28th MAFF Conference of River and Coastal Engineers. University of Loughborough, London.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. Numerical Recipes in Fortran: The Art of Scientific Computing, second ed. Cambridge University Press, Cambridge, UK.

Robinson, A.R., Tomasin, A., Artegiani, A., 1973. Flooding of Venice, phenomenology and prediction of the Adriatic storm surge. Quarterly Journal of the Royal Meteorological Society 99 (422), 688–692.

Talagrand, O., Courtier, P., 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. Part 1: Theory. Quarterly Journal of the Royal Meteorological Society 113, 1311–1328.

ten Brummelhuis, P.G.J., Heemink, A.W., van den Boogaard, H.F.P., 1993. Identification of shallow sea models. International Journal of Numerical Methods Fluids 17, 637–665.

Thacker, W.C., 1988. Fitting models to inadequate data by enforcing spatial and temporal smoothness. Journal of Geophysical Research 93 (C9), 10655–10665.

Thacker, W.C., Long, R.B., 1988. Fitting dynamics to data. Journal of Geophysical Research 93, 1227–1240.

Tomasin, A., 1972. Auto-regressive prediction of sea level in the northern Adriatic. Rivista Italiana di Geofisica XXI, 211–214.

Trigo, I.F., Davies, T.D., 2002. Meteorological conditions associated with sea surges in Venice: a 40 years climatology. International Journal of Climatology 22, 787–803.

Verboom, G.K., de Ronde, J.G., van Dijk, R.P., 1992. A fine grid tidal flow and storm surge model of the North Sea. Continental Shelf Research 12 (2/3), 213–233.

Vested, H.J., Jensen, H.R., Petersen, H.M., Jorgensen, A.M., Machenhauer, B., 1992. An operational hydrographic warning system for the North Sea and the Danish belts. Continental Shelf Research 12 (1), 65–81.

Vested, H.J., Nielsen, J.W., Jensen, H.R., Kristensen, K.B., 1995. Skill assessment of an operational hydrodynamic forecast system for the North Sea and Danish Belts. In: Lynch, D., Davies, A. (Eds.), Quantitative Skill Assessment for Coastal Ocean Models, Coastal and Estuarine Studies, vol. 47. American Geophysical Union, Washington, USA, pp. 373–398.