PROBABILISTIC SEA LEVEL FORECAST FOR VENICE: ENSEMBLE SYSTEM AND “DRESSED” DETERMINISTIC PREDICTION

RICCARDO MEL(1), PIERO LIONELLO(2), LUCA CARNIELLO(3) & LUIGI D’ALPAOS(4)

(1) Dipartimento di Ingegneria Civile, Edile ed Ambientale, Padova, Italy, riccardo.mel@dicea.unipd.it
(2) Dipartimento Scienze e Tecnologie Biologiche e Ambientali, Lecce, Italy, e-mail: piero.lionello@unisalento.it
(3) Dipartimento di Ingegneria Civile, Edile ed Ambientale, Padova, Italy, luca.carniello@dicea.unipd.it
(4) Dipartimento di Ingegneria Civile, Edile ed Ambientale, Padova, Italy, luigi.dalpaos@unipd.it

ABSTRACT
This contribution describes recent work on the development of an operational sea level (SL) prediction system for the city of Venice. The quality of sea level forecast is crucial for the management and maintenance of this city and for operating the movable barriers that are planned to become operational in 2016 for its protection. Recent studies have described the development of an Ensemble Prediction System (EPS) and its effectiveness for delivering a fully informative prediction, including a probabilistic forecast for SL in Venice. This paper discusses a simplified procedure where the full implementation of this method, which requires performing multiple simulations with deterministic models, is avoided. The procedure “dresses” the deterministic prediction of SL in the northern Adriatic Sea with an estimate of its uncertainty that is based on a simple algorithm. This is possible because uncertainty in sea level forecast is to a large extent caused by the uncertainties affecting the forcing meteorological fields, on the predicted value of sea level and it increases with the forecast time range. Here, the results of an operational forecast procedure applied for a 3-month long period in the year 2010, during which an exceptional sequence of storm surges occurred, are shown and compared to Italian and Croatian tide gauges distributed along the coast of the Adriatic Sea. It is shown that the dressed prediction can provide an acceptably realistic estimate of the uncertainty.

Keywords: Ensemble, storm surge, Venice, probabilistic forecast, flood forecast

1. INTRODUCTION
Providing a fully informative sea level (SL) forecast is important in a wide range of problems related to coastal environments protection. Several countries have developed flood forecast systems (e.g. Verlaan et al., 2005; Lane et al., 2009) as an essential tool for efficient managing of coastal defences and delivery of flood warning to alert population. In the last decades, the need to integrate uncertainty estimation into model predictions has been pointed out by many authors (see, e.g., the early contribution of Beven, 1988 and the more recent studies of McMillan, 2008 and Zhong, 2010). Forecast systems of SL need to deliver not only a prediction with the highest possible level of precision, but also information on uncertainty of the forecast and on probability of crossing critical SL thresholds. Clearly, probabilistic forecast and forecast uncertainty are closely linked together and can be produced by the ensemble prediction technique as described in section 2. Techniques for delivering a probabilistic SL forecast have been recently described (Flowerdew et al. 2010, Flowerdew et al. 2012. Mel and Lionello 2014a and 2014b).

The city of Venice is located at the shore of the Northern Adriatic Sea and represents the most important Italian urban centre threatened by storm surges. The dynamics of surges in the Adriatic Sea has been described by several studies (Canestrelli et al., 2001; Lionello, 2005). They are caused by low pressure systems, whose centres are mostly positioned in the north-western Mediterranean Sea, determining a pressure gradient and a south-easterly wind (Sirocco) that blows along the main axis of the Adriatic accumulating water at its northern coast. Venice is an example of a location where an efficient forecast plays a crucial role, especially in relation with the system of movable dams that is presently been built across the three inlets of the Venetian Lagoon. Operating such system requires decision on lifting the barriers to be taken about 8 hours before water reaches 84cm above present mean SL (Eprim et al. 2005). An accurate forecast of SL and on its uncertainty is extremely useful for operating efficiently this movable dam system.

Nowadays in Venice, prediction of SL is delivered by the tide forecast centre (ICPSM - Istituzione Centro Previsione e Segnalazione Maree) of Venice, which adopts a set of linear regression and hydro – dynamical complementary models.
The ICPSM linear regression model is statistically calibrated using observed sea level time series and predicts the water level in the lagoon using observed local SL and observed mean sea level pressure (MSLP) at several Adriatic stations. Hydro- dynamical models integrate the “shallow water” equations and compute the evolution of current and sea level from a sequence of MSLP and surface wind fields. A Hydrodynamic model of ICPSM (the same used in this study and called HYPSE) includes also a data assimilation procedure, where observations available before the initial time of the forecast are used for optimizing the initial condition of the SL forecast (Lionello et al., 2006). Though this set of models has successfully provided SL predictions for many years, it contains unavoidable imprecisions due to inaccurate meteorological predictions, shortcomings of the adopted SL models and errors in the initial condition of the forecasts.

The effectiveness of EPS for delivering an estimate of the forecast uncertainty and of the probability of crossing given SL thresholds has been demonstrated in recent studies (Mel and Lionello 2014a and 2014b). However, the EPS requires to operate multiple runs of the HYPSE model for providing a single probabilistic forecast and needs an adequate operational set-up, including adequate computer resources. This study describes a simplified procedure that can be used for providing an estimate of the uncertainty avoiding multiple simulations. The procedure is based on an algorithm computing the uncertainty of the SL prediction on the basis of the uncertainties affecting the forcing meteorological fields and on the predicted value of sea level. The algorithm, which is calibrated using the EPS results, “dresses” the deterministic prediction of SL in the northern Adriatic Sea with an estimate of its uncertainty as described in section 2.

This contribution first describes the EPS system and the “dressed” deterministic prediction (section 2). Further it summarizes some major results of the EPS in term of representing the forecast uncertainty and briefly compares the skill of the EPS and of the “dressed” deterministic forecast (section 3). Discussion and conclusions are in section 4.

2. DATA, MODEL AND METHOD

The EPS samples the uncertainty inherent in weather prediction to provide more information about possible future weather conditions. The EPS consist on running the meteorological model a number of times from slightly different starting conditions, which are designed to represent small uncertainties inherent in the operational analysis and are specified on the basis of the singular vector technique (Buizza and Palmer, 1995, Molteni et al., 1996). The complete set of forecasts is referred to as the ensemble, and individual forecasts within it as ensemble members. The initial differences between the ensemble members are small, and consistent with uncertainties in the observations. The first implementation in weather forecasting is operational at ECMWF since 1992 and nowadays is a consolidated tool for operational probabilistic weather and storm surge prediction. It is already operationally used by the UK Environment Agency, for issuing coastal flood warnings in England and Wales, and by the Storm Surge Warning Service (SVSD) of Rijkswaterstaat, the Public Works and Water Management Authority in The Netherlands, for the Dutch coast.

The target of this work is to obtain the same information of a EPS forecast using only the deterministic forecast and the meteorological fields uncertainty.

The forecast are computed using HYPSE (Lionello et al., 2006) which computes the SL anomaly produced by wind and sea level pressure, called here sea level residual (SR) and add it linearly to the astronomical tide for obtaining the actual SL. This is justified because of the modest tidal range, which makes nonlinear interactions with the surge negligible. HYPSE is forced by three different sets of 3-hourly ECMWF 10m-wind and MSLP fields: the high resolution meteorological forecast (deterministic forecast, DF) and the 50 ensemble members of the ECMWF EPS, which is a total of 50 + 1 sequences of meteorological fields representing 50 + 1 different evolution of the weather. The implementation of HYPSE used in this contribution does not include the data assimilation component.

The used version of HYPSE adopts a rectangular mesh grid of variable size, which has the minimum grid step (0.03 degrees) in the northern part of the Adriatic Sea, from where grid step increases with a 1.01 factor in both latitude and longitude (in practice, resolution varies in the range from 3.3 to 7km). This grid has been shown to produce more accurate results with respect to other grids (Lionello et al., 2005) and it has been used in a previous study, where advantages of the SR EPS with respect to the deterministic forecast (DF) have been discussed (Mel and Lionello, 2014b). The domain contains a unique open boundary (corresponding to the Otranto Strait connecting the Adriatic Sea to The Ionian Sea) across which a fixed sea level is imposed. This means that the effect of changes of SL in the Ionian Sea on the Adriatic SL are not accounted. Further, this study does account neither for steric effects on SL, nor for changes of total mass of the Adriatic Sea. This limitation is compensated by a bias removal technique (as described in Mel and Lionello 2014a), which add to the SR prediction the effect of long term (several days to month) variability of the mean sea level in the Adriatic Sea. This correction is obtained by subtracting to the forecast the difference between the SR value at the beginning of the forecast and the value obtained by linear interpolation of the observed hourly SR data during the previous day. Figure 1 shows HYPSE domain and the position of tide gauges used in this study.
Forcing resolution of HYPSE is for DF fields T1279 and for ECMWF EPS fields T639. The 10m wind and MSLP fields have been downloaded at 0.125degs (DF) and 0.25degs (EPS) and linearly interpolated to the HYPSE grid (which is the same for all simulations). In the simulated operational prediction stream, the 50 +1 HYPSE runs are carried out twice per day, forced by the ECMWF meteorological fields at 00 and 12UTC. Each HYPSE run provides a 6day forecast and is initialized by a ten day analysis run, where HYPSE is driven by the high resolution (T1279) ECMWF analysis, so that the initial condition and the version of the HYPSE model are the same in all 50 + 1 simulations with the same initial date) and a six day forecast run, each forced by different meteorological forcings, which produce 50 +1 different SR evolutions. The spread of the 50 members around the EMF can be used for an estimate of the forecast uncertainty (Mel and Lionello 2014b). For any selected threshold the fraction of EPS members above it at a given time provides an estimate of the probability of crossing it in the future evolution of the SL (Mel and Lionello 2014b).

Estimation of the forecast uncertainty can be alternatively obtained “dressing” the DF with a prescribed uncertainty computed on the basis of the characteristics of meteorological fields and consequent surges. This method is useful when it is necessary to avoid running 50 EPS members (which requires a considerable computational cost). Since the uncertainty in the sea level forecast mainly stems from the uncertainty affecting the meteorological fields; forecast uncertainty can be estimated via a linear combination of suitable meteorological variances, directly extracted from the meteorological fields. This method has been applied to estimate the uncertainty in the storm surge forecast in the Venice Lagoon (Mel et al. 2014c). The results have shown that the uncertainty estimated through a linear combination of suitable meteorological variances approximates well that obtained using the EPS and can overcome some intrinsic limitations in the use of statistical surge models.

This less computationally expensive method can be used for producing a probability forecast by combining the DF with a prescribed (Gaussian) probability distribution to obtain a probability density function (PDF). This new kind of forecast is called “adjoint extended meteo dressed probabilistic forecast” (ADEMPF). Its Gaussian probability distribution assumes a standard deviation which is the mean of two terms: the first is the sum in quadrature of the half mean of the overall DF rms error and a percentage of the SR predicted by the DF that has been assumed equal to 33%. This first term describes the increasing of uncertainty with the SR level. The second term is a linear combination of the variance four suitable pressure gradients in the Mediterranean Sea. These gradients, usually used in Venice surge forecast models (Canestrelli 2004), are computed between: CNR platform and Bari, CNR platform and Zara, CNR platform and Genova and between Genova and Porto Torres. These gradients can describe the meteorological input uncertainty.

The idea of obtaining a probabilistic forecast (PF) by associating to a DF a parametrized estimate of its uncertainty has already been used in the literature (Flowedew et al. 2012). Here the adjoint extended meteo dressed probabilistic forecast is adopted. It assumes a standard deviation that is the mean of two terms: the first is the sum in quadrature of the half mean of the overall DF rms error and a percentage of the SR predicted by the DF that has been assumed equal to 33%. This is meant to describe the increase of uncertainty with the SR level. The second term is a linear combination of the variance four suitable pressure gradients in the Mediterranean Sea. These gradients, usually used in Venice surge forecast models (Canestrelli 2004), are computed between: CNR platform and Bari, CNR platform and Zara, CNR platform and Genova and between Genova and Porto Torres.
3. RESULTS

The EPS is the ideal tool for proving PF of a predefined event. Given the set of a threshold level $h$, the event $H$ is defined as the SR exceeding $h$ for at least one hourly step within a 12h interval, where the a set of 12 intervals is specified covering the whole time range from the beginning to the end of each forecast, which is 144h long. On the basis of the EPS the probability $p(h)$ is computed as the fraction of EPS members in which the event $H$ occurs. EPS statistics include SR EPS spread, SR forecast rms error with respect to tide gauges observations and the Brier Skill Score (BS, Brier 1950) to measure the accuracy of the probability estimate by the EPS. It is expected that the EPS spread is indeed large when the EMF and DF errors are large and it represents a good estimate of the uncertainty of the prediction.

Fig.2 summarizes the EPS forecast features by aggregated data over all five tide gauges and considering all forecast lead times by comparing DF and EMF performance. The highest SR forecast values have the largest uncertainty and are most likely to be appreciably wrong (Fig 2a). However, percent-wise the forecast errors and the EPS spread as well, are smaller for large than for small SR anomalies. For large SR EMF anomalies EMF rms error is lower than DF rms error, showing a clear improvement and robustness of the EMF with respect to the traditional techniques.

![Figure 2](image)

Figure 2. Left panel (a): relation between Ensemble Mean Forecast magnitude and rms error of EMF (red line with triangles), DF (green line with squares) and all Ensemble members (ENS, grey line). Right panel (b): rms errors as function of the RS-EPS spread for Ensemble Mean Forecast (EMF, red line with triangles), Deterministic Forecast (DF, green line with squares) and all Ensemble members (ENS, grey thick line).

Fig. 2b shows the relation between the EPS spread and the rms error. The EPS spread is linked to the EMF rms error except for very small errors. This is because in the initial period of the forecast EMF errors are small, but not nil, while the corresponding EPS spread is zero. In other words, a small spread cannot be associated to a vanishing error because the errors in the initial state. On the contrary the EPS spread increases faster than the rms errors, so that the largest errors of EMF (and the DF as well) are appreciably smaller than the corresponding EPS spread. This suggest that the EPS overestimates the uncertainty associated with large departure of individual ensemble members from the EMF.

The Brier score and Brier skill score, commonly used in measures for evaluating probability forecasts (Roulston, 2007), are computed on the basis of the EPS as function of the lead time for two different thresholds $h$ (Fig. 3). Data are aggregated considering all tide gauges. Results indicates that EPS BSS remains positive for very large lead times and diminishes with the threshold (red thick line). Therefore the skill of the PF decreases for increasing thresholds, but it remains positive for 4 day seven for the highest considered threshold (50cm).

In this paper three different kinds of estimating probability are compared. The first is an undressed probability forecast (UPDF) that assumes a perfect forecast with no uncertainty so that the probability of $H$ is 1.0 if it occurs in the DF, 0 otherwise (this can be considered a limit case in which a Gaussian with zero standard deviation is adopted). The second is a mean dressed forecast (MDPF) that assumes a PDF centred on the DF with a prescribed standard deviation which is an estimation of the rms error of the DF. Finally the ADEMDPF, as described in section 2, is also shown.

The BSS has been computed using the EPS results and three simple methods (UDPF, MDPF, ADEMDPF) for producing the PF. Fig.3 shows the BSS for the thresholds $h=20$cm and $h=50$cm. The BSS of the EPF systematically outperforms those of the other PF methods. Particularly the BSS of the UDPF, which does not account for any error in the forecast, performs poorly. In fig 3a, even for $h=20$, value for which the BSS of all other methods remain consistently above 0.2 for a 3-day long period, the BSS of UDPF drops below 0.2 already after 30 hours. Considering the $h=50$cm threshold the UDPF BSS drops below 0 after 48 hours. This is a strong indication that a probability estimate based on a plain DF approach cannot deliver reliable results. The ADEMDPF, has a high BSS and could be used as a cheap alternative to the much more expensive EPF.
4. CONCLUSIONS

This contribution describes how to deliver a probabilistic forecast for the city of Venice. Increased requirements in the information to be delivered by a SL prediction system associated to sophisticated operational requests have stimulated innovative techniques. The prediction of SL in Venice is, since many years, regularly carried out for effective warning of the population and reducing damages produced by severe events. However, recent requests for a more informative and accurate system, mainly associated with operating the movable dams (MOSE), are requiring a probabilistic prediction.

The EPS is the ideal tool for providing a probabilistic prediction. Studies have analysed (Flowerdew et al., 2010; Flowerdew et al., 2012; Mel and Lionello 2014a and 2014b) the quality of the information delivered by the EPS and its advantages with respect to the more common DF. EPS provides information by calculating the probability of exceeding a specified critical level during the forecast time window. Moreover it is possible to use EPS for computing the uncertainty in the timing of the peak and in general of the SL forecast: SR peak could be predicted with good accuracy and the EPS spread would have allowed delivering a warning the occurrence of SR levels above critical thresholds several days in advance. Generally, the EMF has rms error lower than DF, especially for short (up to 3 days) lead times. Though differences are not large, EMF produces a more robust prediction of the peak values of SR (Mel et al., 2014a) than the DF.

The EPF has a clear skill in predicting the actual probability distribution of SR, while a probability estimate based on a single DF is shown to be inadequate. However a probabilistic forecast obtained with a prescribed Gaussian centred on the DF value accounting for the dependence of the DF rms on the predicted SR values and lead times performs very similarly to the EPS based PF. Therefore, ADEMDPF can be considered a practical computationally cheap alternative to EPF. This method can be applied in other hydraulic problems, using the correlation between the output uncertainty and both input meteorological data and the forecast features, and it is very useful when avoiding the computational costs of multiple runs of hydrodynamic models is needed.
ACKNOWLEDGMENTS

The hourly SL observations in CNR-PLATFORM and Venice “Punta Salute” were provided by I.C.P.S.M. of the Venice City Council, observations in Trieste by Dr. Fabio Raichich of ISMAR-CNR (Italy), observations in Rovinj, Split, Dubrovnik by Dr. Nenad Leder of the Croatian Hydrographic Institute. Founding has been provided by 2013 University of Padova project “Combined use of remote sensing and in situ measurements for the calibration of transport and diffusion models in shallow coastal lagoons” (CPDA133253/13).

REFERENCES