DATA ASSIMILATION STUDY USING THE DUTCH CONTINENTAL SHELF MODEL WITH FULL MEASUREMENT

JULIUS H. SUMIHAR[1] AND MARTIN VERLAAN[1,2]

[1]Faculty of Electrical Engineering, Mathematics and Computer Science, Department of Applied Mathematics, Delft University of Technology, Delft, The Netherlands
[2]Dutch National Institute for Coastal and Marine Management (Rijkwaterstaat/RIKZ), Den Haag, The Netherlands

Abstract

The operational sea water level prediction system in the Netherlands is based on the decomposition of water level into the astronomical tides and the surge. While the astronomical tides are analysed and predicted by using Harmonic Analysis, the surge is predicted by using numerical hydrodynamics model named Dutch Continental Shelf Model, DCSM. Using this approach the nonlinear interaction between the two components is not well accomodated. Moreover, the performance of the system based on this decomposition now seems to have reached its limit. In attempt to further improve the prediction we are going to apply data assimilation using the DCSM model with the full water level measurements without any decomposition. In the first step of the study we use the steady-state Kalman filter as the method for data assimilation. As the success of data assimilation depends largely on the error representation, we also work with a new error representation for the open boundary condition. In the new error representation, a colored noise process, modelled using AR(1), is assigned to each harmonic parameters defining the water level boundary conditions. This choice is made based on the fact that the harmonic parameters of the astronomical tides are in fact slowly varying in time.

1. Introduction

Accurate sea water level forecasting is important in the Netherlands. This is mainly because large areas of the land lie below sea level. Forecast are made to support storm surge flood warning system. Accurate forecast of at least 6 hour ahead are also needed for proper closure of the movable storm surge barriers in the Eastern Scheldt and the New Waterway. Moreover, it is also important for harbour management, as the size of some ships have become large that they could only enter the harbour during high water period.

The operational sea water level prediction system in the Netherlands is based on the decomposition of water level into two components: the astronomical tides and the surge. Each of these components is analysed and forcast separately. While the astronomical tides are analysed and predicted by using Harmonic Analysis [*Godin*,1972], the surge is predicted by using numerical hydrodynamics model named Dutch Continental Shelf Model (DCSM) [*Stelling*, 1984; *Verboom et al.*,1992]. Hundreds of tidal constituents are used in the harmonic analysis. On the other hand, the numerical model DCSM has been through some development as well. Among the latest developments is the increased resolution of the wind forecast input from 55 to 22 km, which could slightly increase the forecast quality in storm situations. Another attempt was to refine further the resolution of the wind input to 11 km. However case studies comparing resolution 22 km and 11 km input for DCSM showed that the 11 km resolution could not improve the prediction. Besides works on improving the wind input, a lot of works had also been devoted to model calibration. The last calibration was performed in 1998. It is now assumed that, with the limited model resolutions of approximately 8 km and present quality of the bathymetry information, further calibration is not worthwhile. For further details of the development see *Verlaan, et al.* [2005].

Although water level prediction based on this decomposition is a common practice, it has some difficulties. Using this approach the nonlinear interaction between the components is not well accomodated. It is standard practice to remove harmonic tides from measure water levels and then model the remaining residual. Because of the magnitude of astronomical forcing and the limitations of analytic tools, analysis of water levels has traditionally emphasized linear methods to decompose water levels into tides and other components. This decomposition presupposes that water levels are a superposition of independent linear components, despite the fact that the inherently nonlinear Navier-Stokes equations are used to describe these motions. *Frison* [2000] showed that the dynamical characteristics of the harmonic model and the dynamical characteristics of the residuals have no relationship to the dynamical characteristics of the observed water levels. This finding suggests that the traditional linear decomposition of water levels may not be an appropriate means of modeling water levels in coastal regions and estuaries.

As an attempt to further improve the prediction we are going to apply data assimilation using the DCSM model with the full water level measurements without any decomposition. Following this approach is expected to be able to better represent the tide-surge interaction, which in turn to lead to better forecast accuracy. In the first step of the study we use the steady-state Kalman filter as the method for data assimilation [*Heemink*, 1986].

It is known that the success of data assimilation depends largely on the error representation. In this study we will apply a new error representation for the boundary condition. In the new representation, a colored noise process, modelled using AR(1), is assigned to each harmonic parameters defining the water level boundary conditions. This choice is made based on the fact that due to its interaction with other phenomena the harmonic parameters of the astronomical tides are in fact slowly varying in time [*Heemink et al.*,1991].

This paper is organized as follows. In Section 2 a brief description about the deterministic hydrodynamics model of the DCSM is presented. This is followed by the description of the operational uncertainty model in Section 3 and the new uncertainty representation for the open boundary condition in Section 4. Section 5 presents and discusses some results of the study, while Section 6 give the conclusion.

2. Deterministic Model Equations

The hydrodynamic equations system used in the DCSM is the 2-D shallow water equations, which is based on the conservation of mass and momentum [*Stelling*, 1984; *Verboom* *et al.*,1992]. The conservation of mass and momentum are expressed as

$$\frac{\partial \zeta}{\partial t}+\frac{\partial (Hu)}{\partial x}+\frac{\partial (Hv)}{\partial y}=0 \qquad (1)$$

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+g\frac{\partial \zeta}{\partial x}-fv+\frac{gu\sqrt{u^2+v^2}}{C^2H}-C_d\frac{\rho_a}{\rho_w}\frac{V^2cos\psi}{H}+\frac{1}{\rho_w}\frac{\partial p_a}{\partial x}=0 \qquad (2)$$

$$\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+g\frac{\partial \zeta}{\partial y}+fu+\frac{gu\sqrt{u^2+v^2}}{C^2H}-C_d\frac{\rho_a}{\rho_w}\frac{V^2sin\psi}{H}+\frac{1}{\rho_w}\frac{\partial p_a}{\partial y}=0 \qquad (3)$$

where (x,y) the cartesian coordinates in horizontal plane, u and v are the depth-averaged current in x and y directions respectively, ζ and D water level and water depth respectively above the reference plane, $H=\zeta+D$ total water depth, g gravity acceleration, f coefficient for the Corriolis force, C Chezy coefficient, V the wind velocity, ψ wind angle with respect to the positive x-axis, C_d wind friction coefficient, p_a air pressure at the surface, ρ_w density of sea water, and ρ_a density of air at the surface.

These equations are discretized using an Alternating Directions Implicit (ADI) method and a staggered grid that is based on the method by *Leenderstse* [1967] and *Stelling* [1984]. At land-water boundaries the normal flow is set to zero, while water levels are specified at the open boundaries using ten harmonic constituents. As the wind input, this model uses forecasts of the meteorological high-resolution limited area model (HiRLAM), provided by the Netherlands Meteorological Institute (KNMI). This hydrodynamic numerical model was made available by means of the WAQUA package, a software system for 2D shallow water simulation of the Dutch National Institute for Coastal and Marine Management (Rijkwaterstaat/RIKZ).

3. Uncertainty Models

Before a Kalman filter can be applied a description of the errors in the model and the measurements are needed. An important tool for modelling the errors as a random or stochastic process in the model as well as measurement is ARMA process [*Box et al.*, 1994]. An ARMA model maps a white noise process to an error process with the desired shape of the autocorrelation function. It is assumed that the main sources of error in the DCSM are the open boundary condition and the wind input as provided by the meteorological model. The covariances of these errors are modelled using ARMA models.

3.1. **Uncertainty model at the open boundary.** While in general it is possible to use other variables, the DCSM uses water level as the specification for its open boundary condition. The water level at the open boundaries is determined by specifying amplitudes and phases of tidal constituents [e.g. *Godin*, 1972]. The DCSM uses ten tidal constituents (M2, S2, N2, K2, O1, K1, Q1, P1, NU2, L2) for defining the open boundaries. In a simplified format the open boundary condition reads

$$\zeta_b(k)=\Sigma_i A_i cos(\omega_i t_k)+B_i sin(\omega_i t_k) \qquad (4)$$

where A_i and B_i are the harmonic parameters of tidal constituent i.

It is common practice to specify the amplitudes and phases of the components only once. However, due to other process, e.g. meteorological effect, the water level differs somewhat from what is predicted using harmonic analysis. The change of the water level due to meteorological effects remains unknown. Here the meteorological set-up at the open boundary is modelled as an AR(1) process in time:

$$\Delta\zeta_b(k) = \alpha_b \Delta\zeta_b(k-1) + \epsilon_b(k) \tag{5}$$

where $\epsilon_b(.)$ is a zero mean white noise process. The set-up at the open boundary is only specified on a few grid points. The set-up at the other points of the open boundary are computed using linear interpolation.

3.2. **Uncertainty model for the wind input.** The wind forcing, obtained from meteorological forecasts, is still a major source of error for water level forecasts using shallow water models. Besides the wind forecast input, the model that translates them into stresses on the water surface is also not accurate. In addition, the coefficient C_d in the model is not known accurately, nor is it believed to be really constant [*Verlaan*, 1998]. All these effects cause an error in the surface stress that is used in the model.

The uncertainty of the meteorological forcing is represented as a noise process, which is added to the depth averaged momentum equations. This process is constructed using an AR(1) process for the correlation in time:

$$\Delta u(x,y,k) = \alpha_w \Delta u(x,y,k-1) + \epsilon_u(x,y,k) \tag{6}$$

$$\Delta v(x,y,k) = \alpha_w \Delta v(x,y,k-1) + \epsilon_v(x,y,k) \tag{7}$$

while the spatial characteristic is specified by using spatial correlation given by

$$\frac{E[\Delta u(x_1,y_1,k)\Delta u(x_2,y_2,k)]}{\sqrt{E[\Delta u(x_1,y_1,k)^2]E[\Delta u(x_2,y_2,k)^2]}} = e^{-\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}/d_w} \tag{8}$$

$$\frac{E[\Delta v(x_1,y_1,k)\Delta v(x_2,y_2,k)]}{\sqrt{E[\Delta v(x_1,y_1,k)^2]E[\Delta v(x_2,y_2,k)^2]}} = e^{-\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}/d_w} \tag{9}$$

$$\frac{E[\Delta u(x_1,y_1,k)\Delta v(x_2,y_2,k)]}{\sqrt{E[\Delta u(x_1,y_1,k)^2]E[\Delta v(x_2,y_2,k)^2]}} = 0 \tag{10}$$

where

$$\sqrt{E[\Delta u(x_1,y_1,k)^2]} = \sqrt{E[\Delta v(x_2,y_2,k)^2]} = \sigma_w \tag{11}$$

and d_w is a characteristic distance for the correlation. The noise is introduced on a coarser grid to reduce the number of noise inputs. The noise at other grid points are interpolated subsequently by using bilinear interpolation.

4. New representation of uncertainty at open boundary

The success of a data assimilation scheme depends largely on the error representation. Therefore, as a further attempt of improving water level forecast, we also performed a study of data assimilation scheme using a new representation of error at the open boundary condition.

It is known that the astronomical tides interact with other phenomena, e.g. surge. This interaction is manifested in term of a number of narrow bands centered at tidal frequencies in the spectrum of water level residual [*Heemink et al.*, 1991; *Pugh*, 1987]. Residual is defined as observed water level minus astronomical tides as predicted using harmonic analysis. It can be shown that a narrow band process can be modelled as a sinusoidal function with slowly varying amplitude and phase, whose frequency is constant and equal to the center-frequency of the narrow band [*Heemink et al.*,1991]. To accomodate this slowly varying behaviour, the tidal equation is now written as:

$$\zeta_b(k) = \Sigma_i (A_i + \Delta A_i(k)) cos(\omega_i t_k) + (B_i + \Delta B_i(k)) sin(\omega_i t_k) \tag{12}$$

where $\Delta A_i(k)$ and $\Delta B_i(k)$ represent the slowly varying component of the harmonic parameter of tidal constituent i. Rearranging the equation, the meteorological set-up is now given by

$$\Delta\zeta_b(k) = \Sigma_i \Delta A_i(k) cos(\omega_i t_k) + \Delta B_i(k) sin(\omega_i t_k) \tag{13}$$

Furthermore, *Heemink et al.* [1991] also showed that the variation of the harmonic parameters can be modelled by means of the stochastic differential equation, whose discrete approximation takes the form of AR(1) process as follows:

$$\Delta A_i(k) = \alpha_A \Delta A_i(k-1) + \epsilon_A(k-1) \tag{14}$$

$$\Delta B_i(k) = \alpha_B \Delta B_i(k-1) + \epsilon_B(k-1) \tag{15}$$

where $\epsilon_A(.)$ and $\epsilon_B(.)$ are zero mean white noise processes. As in the operational set-up, the set-up at the open boundary is only specified on a few grid points, while the set-up at the other points are computed using linear interpolation.

5. Results and discussion

In the operational system, the Kalman filter is executed every six hours to provide the best initial condition for the next water level prediction. In this study, we evaluate the performance of three different schemes in term of the accuracy of the initial conditions estimated by each scheme. These schemes are the deterministic model without data assimilation, the scheme with data assimilation using original boundary error representation (Equation 5), and the one with data assimilation using the new boundary error representation (Equation 13).

In this study, water level field data of observation period November 1, 2004 - October 1, 2005 is used for the experiments. As in the operational setting, eight different stations are used as the assimilation stations. The locations of these stations are depicted in Figure 1. This selection was based on the assumption that the storm surge dynamics were dominated by the propagation of the tidal Kelvin wave down the British coast together with the direct stress exerted by the wind on the water surface [*Verlaan, et al.*, 2005]. Furthermore, measurements for validation are available for eight different stations, of which one is located on the east coast of England, six along the Dutch coast, and one on southern North Sea. For the wind input, analysed wind is used instead of forecast wind. Hence the experiment accords the operational setting, except for the fact that the operational setting uses forecast wind.

For the scheme with original boundary noise representation, the filter parameters are set the same as the operational values. At the open boundary, the standard deviation of the water level is 6.8 cm with the time correlation of $\alpha_b = 0.9$ per time step. This set of parameters was found by manual calibration. Another effort of calibration using covariance matching confirmed these values [*Verlaan*, 1998]. On the other hand, for the scheme with new boundary error representation, the standard deviation of each of the noise processes $\Delta A_i(k)$ and $\Delta B_i(k)$ was found by using manual calibration to be 0.23 cm. Noise processes of Equation 5 and 13 are located at the five corners of the open

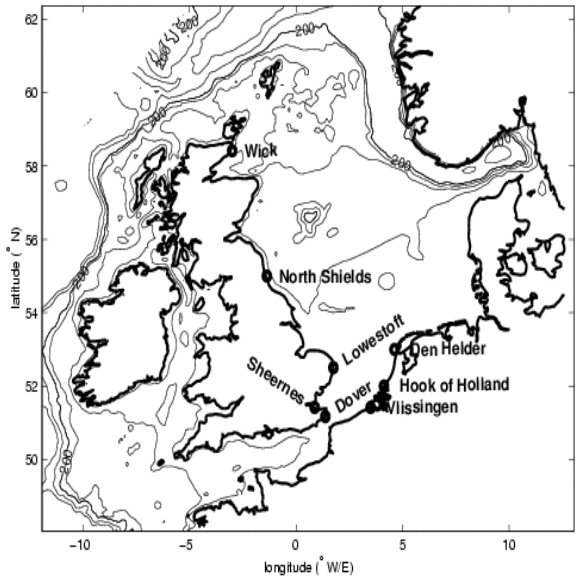

FIGURE 1. DCSM area and assimilation stations

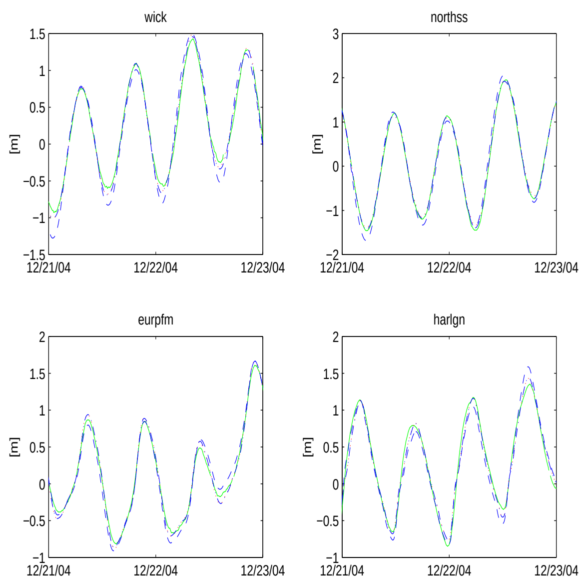

FIGURE 2. Water-level timeseries at Wick, North Shields, Europlatform, and Harlingen: observation (solid), deterministic (dashed), original scheme (dotted), new scheme (dashdot)

boundary for this purpose. Furthermore the wind noise is introduced on a coarse grid with increments of 28 in both directions. A standard deviation of 0.0068 m/s per time step is used together with a time correlation of $\alpha_w = 0.9$ and a spatial correlation length of 26.38 grid cells. Furthermore, the standard deviation of the measurement noise is set to 10 cm. Figure 2 presents some of the experiment results using these schemes for two assimilation stations and two validation stations.

Table 1. Root mean square residual at assimilation stations (in centimetres)

Station	Deterministic	Original Scheme	New Scheme
Wick	14.8	6.5	6.5
North Shields	14.2	7.6	7.0
Lowestoft	14.8	5.6	5.5
Sheerness	34.6	23.3	22.6
Dover	17.6	11.1	13.8
Vlissingen	18.0	12.5	12.6
Hoek van Holand	14.1	10.1	9.2
Den Helder	16.0	11.2	11.0
Mean	18.0	11.0	11.0

Table 2. Root mean square residual at validation stations (in centimetres)

Station	Deterministic	Original Scheme	New Scheme
Aberdeen	12.0	9.3	8.8
Europlatform	11.2	5.6	5.1
Cadzand	18.3	15.6	13.4
Roompot buiten haven	18.6	13.7	12.8
Ijmuiden buiten haven	14.9	9.8	9.1
Harlingen	15.1	11.3	11.4
Huibertgat	16.2	12.9	13.3
Delfzijl	27.7	26.8	26.9
Mean	16.7	13.1	12.6

As a measure for comparing the performance of the different schemes we use the root mean square residual, where residual is defined by observed water level minus water level as estimated by each scheme. This measure gives a clear indication of how close the estimated water level to the observed one. The evaluation was done in both the assimilation and validation stations. The results obtained for the assimilation and validation stations are presented seperately in Table 1 and 2. The three columns of each table give the rms residual as estimated by using the deterministic, data assimilation with original boundary noise, and data assimilation with new boundary noise schemes respectively. To provide a general idea about the performance, the last line of each table give the mean of data at each column.

From these tables we see that applying data assimilation improves the estimate obtained by using deterministic model. The tables also indicate some improvement obtained from implementing new boundary error representation at some individual stations. The improvement varies from station to station. On the other hand, at a few other individual stations the original scheme performs better than the new one. However, Table 2 shows that the overall performance of the new scheme slightly surpasses the original one. With respect to the performance of the deterministic model at individual station, the overall average of improvement obtained from using the old and new schemes is 32.5% and 34.1% respectively.

6. Conclusion

As an effort to improve the accuracy of water level prediction, a study of implementing data assimilation using DCSM with full measurement was performed. Without decomposing water level into astronomical tide and surge, it is expected to accomodate better the tide-surge interaction, thus yield better forecast accuracy. A steady-state Kalman filter was used for this purpose. Moreover, a new boundary error representation was also implemented. Here, two noise processes, modelled as AR(1) processes, are defined for each harmonic constituent specified at the open boundary.

In this study the performance of three different schemes were compared: deterministic model, Kalman filter with the original boundary noise representation, and Kalman filter with the new boundary noise representation. Experiments were performed using field data of water level, but with analysed wind. Using these input, the output of the simulations is in fact the initial conditions for the next water level forecast in the operational system. The comparison of the performance was done in term of the accuracy of the initial condition estimated by each scheme. Using eight assimilation stations and eight validation stations, the experiments showed that there is only a slight difference between the performance of the original and the new schemes. Improvement of 32.5% and 34.1% were obtained using the original and new schemes respectively.

For the next step, another evaluation of the original scheme based on the forecast performance is necessary, since the system is intended for making forecast of water level. For this purpose forecast wind should be used instead of analysed wind. In addition, a study of selecting assimilation stations may also be necessary to obtain a better result of data assimilation.

References

Box et al. 1994. Box,G.E.P., G.M. Jenkins, and G.C. Reinsel (1994), *Time Series Analyis*, Prentice Hall, Englewood Cliffs, 3 edition

Frison, 2000. Frison,T.W. (2000), Dynamics of the residuals in estuary water levels, *Physics and Chemistry of the Earth (B), 25*, no.4, 359 - 364

Godin, 1972. Godin,G. (1972), *The Analysis of Tides*, Liverpool University Press

Heemink, A.W., 1986. Heemink,A.W. (1986), Storm surge prediction using Kalman filtering, *PhD thesis*, Twente University of Technology, Rijkwaterstaat Communications no. 46.

Heemink et al. 1991. , Heemink, A.W., B. de Jong, and H. Prins, Adaptive harmonic analysis, *Deutsche Hydrographische Zeitschrift, 44, No.2*, 91 - 106

Leendertse, 1967. Lendeertse, J.J. (1967), Aspects of a computational model for long-period free-surface flow modelling, *PhD thesis*, Rand Corporation, RM-5294-RR

Pugh 1987. Pugh,D.T. (1987), *Tides, surge, and mean sea level*, John Wiley & Sons Ltd, Chichester

Stelling, 1984. Stelling, G.S. (1984), On the construction of computational methods for shallow water flow problems, *PhD thesis*, Deflt University of Technology.

Verboom et al., 1992. Verboom, G.K., J.G. de Ronde, and R.P. van Dijk (1992), A fine grid tidal flow and storm surge model of the North Sea, *Cont. Shelf Res. 12*, 213 - 233

Verlaan, 1998. Verlaan,M. (1998), Efficient Kalman Filtering Algorithms for Hydrodynamics models, *PhD thesis*, Delft University of Technology

Verlaan et al., 2005. Verlaan,M., A. Zijderveld, H. de Vries, and J. Kroos (2005), Operational storm surge forecasting in the Netherlands: developments in the last decade, *Phil. Trans. R. Soc. A, 363*, 1441 - 1453.