IJE TRANSACTIONS A: Basics Vol. 30, No. 1, (January 2017) 23-29

International Journal of Engineering

Journal Homepage: www.ije.ir

TECHNICAL NOTE

Enhanced Predictions of Tides and Surges through Data Assimilation

R. R. Karri*[a], V. Babovic[b]

[a] *Petroleum and Chemical Engineering, Universiti Teknologi Brunei, Brunei Darussalam*
[b] *Department of Civil and Environmental Engineering, National University of Singapore, Singapore*

PAPER INFO

Paper history:
Received 29 September 2016
Received in revised form 04 November 2016
Accepted 11 November 2016

Keywords:
Ensemble
Kalman filter
Data model integration
Data assimilation
OpenDA
Tides and surges

ABSTRACT

The regional waters in Singapore Strait are characterized by complex hydrodynamic phenomena as a result of the combined effect of three large water bodies viz. the South China Sea, the Andaman Sea, and the Java Sea. This leads to anomalies in water levels and generates residual currents. Numerical hydrodynamic models are generally used for predicting water levels in the ocean and seas. But their correctness is typically limited by several factors, namely the complexity associated with the coastal geometry and uncertainty in the flow forcing factors like (winds, pressure and deep ocean tides). Modeling of ocean dynamics in the Malacca strait and Singapore regional waters is particularly challenging due to the presence of large number of smaller islands and strongly nonlinear tidal interactions. The complexity is further enhanced due to the composite local bathymetry and geometry variations around the Singapore Island and meteorological effects on different scales. This study acknowledges the enhancement and better prediction of tides and surges through the use of data assimilation. Through a portable interface OpenDA, an ensemble Kalman filter is integrated with a hydrodynamic model to enhance the model predictions. To assess the sensitivity and evaluate model enhancement, a twin experiment is designed to improve tidal boundary forcing effect in a semi-enclosed estuary. The key outcomes of this study signify that the model results can be significantly improved in this complex flow regime.

***doi**: 10.5829/idosi.ije.2017.30.01a.04*

1. INTRODUCTION

Variations in sea levels are mainly caused by oceanographic and meteorological parameters such as winds, sea surface temperature, and prevailing atmospheric pressure. To understand and ascertain the uncertainties associated with the large water bodies, numerical modelling has been found to be a powerful tool to study the variations and contributed factors. Numerical sea models are capable of reenacting the space-time advancement of the sea state, with the resolution important to reproduce the little scale forms that are not captured by the observations, which are usually meager in space and restricted in time [1-3]. Numerical modeling of ocean dynamics in the Malacca strait and Singapore regional waters is particularly challenging due to the presence of a large number of small islands and strongly nonlinear tidal interactions [4]. The complexity is further enhanced due to the composite local bathymetry and geometry variations around the Singapore Island and the meteorological effects at different scales. Flows in this region encounter the impacts of nonlinear dynamical associations between three major water bodies - South China Sea to the east, Andaman Sea to the west and Java Sea to the south. The intricate shallow water hydrodynamics resulting from the various ocean currents moving into and out of this region combined with sharply varying transient meteorological impacts prompt high variability in water levels and currents along the Singapore coast. Given Singapore's pivotal role in shipping in the region, an understanding of the characteristics and precise modelling of such complex phenomena and their effect on the seaside morphology are of great monetary and ecological significance. For better understanding of the processes and reproducing the ocean dynamics through a flow model with high resolution around Singapore

---

*Corresponding Author's Email: kramarao.iitd@gmail.com (R. R. Karri)

Please cite this article as: R. R. Karri, V. Babovic, Enhanced Predictions of Tides and Surges through Data Assimilation, International Journal of Engineering (IJE), TRANSACTIONS A: Basics Vol. 30, No. 1, (January 2017) 23-29

straits, the Singapore Regional Model (SRM) has been created inside the Delft3D numerical demonstrating framework. This integrated model reproduces barotropic streams driven by tides and meteorological compelling.

Regardless of recent enhancements, numerical ocean models still result in deviations in their predictions due to the inadequate representation of sea flow and thermodynamics. This may be because of limited insight into physical mechanisms, simplifying assumptions in numerical implementation, approximation to governing equations and errors in observations. Thus it is important to correct (improve) such models, and this should be possible through data assimilation (DA) [5, 6]. Data assimilation methods synthesize the model arrangement with accessible observations to acquire an optimal solution which can be utilized as the new starting condition for model forecasting. All in all, the analysis procedure minimizes the misfit between the model states and the observations utilizing a least squares strategy or auto regressive models. Kalman filter (KF) based data assimilation is found to be competent in handling high dimensional systems like ocean dynamics and weather forecasting. KF is not only efficient in minimizing the variance of error, but can also improve the solution in areas where there are no observations. The traditional Kalman filter [7, 8] is designed for linear systems and the extended Kalman filter (EKF) extends its basic calculation to non-linear problems by linearizing the non-linear function around the present assessment. For a realistic ocean model, it is hard to apply straightforwardly the EKF because ocean dynamics are unequivocally nonlinear and any linearization fizzles out rapidly.

In this research study, we utilize ensemble Kalman filter (EnKF), which has been previously applied to various non-linear complex systems and is considered as one of the most advanced sequential assimilation methods. This assimilation method is one of the filters available in the portable data assimilation toolbox (OpenDA) for enabling data assimilation and calibrations tools that can be used with arbitrary flow models. OpenDA makes utilization of an arrangement of interfaces that depict the communication between models, observations and data assimilation methods and is available as open source (www. openda.org). OpenDA provides generic data assimilation environment [9] which has been widely applied to a range of physical process models and hydrodynamic unsteady flow models. This environment also features filtering techniques like Ensemble Kalman Filter (EnKF), Ensemble Square Root Filter (EnSQRTF), Reduced Rank Square Root Filter (RRSQRTF), and particle filters [10, 11]. These functionalities have been successfully applied in different areas such as data assimilation of currents and salinity profiles [12], flood forecasting [13] and accurate forecasting of sea level anomalies (SLA) and residual currents [14, 15]. These diverse applications demonstrate the efficacy of OpenDA as a generic tool box for DA. In the present research study, the DA scheme in the framework of EnKF available in OpenDA is used to improve the prediction of water levels and surges in the Singapore straits.

2. DESCRIPTION OF NUMERICAL MODEL

Singapore Regional model (SRM) developed by Kurniawan et al. [4] was used as the base numerical hydrodynamic model. This model was presently setup to provide hydrodynamic information, both tide and sea level anomalies in the Singapore locale. The model was developed using Delft 3D-FLOW framework, which assumes Boussinesq conditions and incorporates the Navier-Stokes equations applicable for shallow water, to simulate the free surface flow for the whole water body between the 3 major water bodies surrounding the Singapore straits. The scope, grid, bathymetry and 3 open boundaries of SRM are shown in Figure 1. The SRM model has a total of 38,500 grid cells with a resolution tailored to the regions and spatial scales of interest; varying from 15 km at the open sea boundaries to around 100 m nearby to Singapore straits. The bathymetry in the SRM is based on Admiralty charts and varies from about 2000 m in the Andaman Sea to less than 5 m in coastal areas around Singapore. The rectangular zone represents a half open simplification of Malacca Strait, which will be discussed in detail in section 5.

3. ENSEMBLE KALMAN FILTER ALGORITHM

The ensemble Kalman filter (EnKF) is a sophisticated sequential data assimilation method introduced by Evensen [16]. It applies an ensemble of model states to represent the error (deviations) measurements of the model estimate and then it applies ensemble integrations to predict the error characteristics forward in time. Finally, it uses an analysis scheme which works straightforwardly on the ensemble of model states, when observations are assimilated.

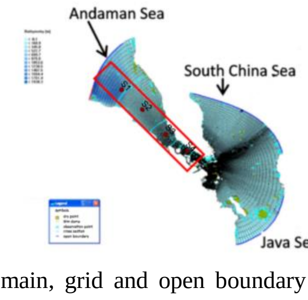


Figure 1. Domain, grid and open boundary transects of the SRM

The traditional Kalman filter depends on the assumption that the likelihood distributions follow a Gaussian distribution, and gives mathematical equations that express the changes in mean and covariance by the Bayesian update, and provides a formula for propagating the covariance matrix in time domain as long as the system is linear. Nonetheless, computationally this is unrealistic for high dimensional systems. To overcome this drawback, EnKF technique was developed [16-18]. EnKF computes the propagation of the system state using an ensemble of simulations with random disturbances.

One key benefit of EnKF is that advancing the likelihood distribution in the time domain is accomplished basically by advancing each individual in the ensemble. The EnKF has been demonstrated to effectively handle strongly nonlinear elements and vast state spaces, and is currently utilized as a part of several practical applications with primitive condition models for the ocean and atmosphere [15, 17, 19, 20]. In this study we consider an algorithm based on the Bayesian update step for the EnKF version in which the filter involves randomization of data.

4. FORMULATION OF EnKF

In the implementation of EnKF, the measurable properties of the state vector are characterized by an ensemble of probable state vectors as shown in Figure 2.

The forecast ensemble consists of *N* members, which are state vectors of dimension *n*. The ensemble can be written as the *N* by *n* matrix:

$$X^f = [x_1, x_2, ..., x_N] = [x_i] \tag{1}$$

The mean of each ensemble is given as:

$$E(X) = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{2}$$

the covariance matrix of each ensemble is:

$$C = \frac{AA^T}{N-1} \tag{3}$$

where

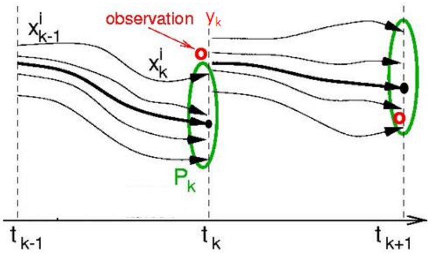


Figure 2. State update and forecast by EnKF

$$A = X - E(X) \tag{4}$$

The measurement data is of size *m* and corresponding error covariance matrix *R* size is *m * m*. Then the measurement matrix *D* (*m*N)* is defined as:

$$D = [d_1, d_2, ..., d_N],\ \mathrm{d_j} = d + v_j; v_j \sim \mathrm{N(0,R)} \tag{5}$$

where v_j are non-correlated random perturbations. Then analysis ensemble is calculated as:

$$X^a = X + \frac{CH'}{(HCH' + R)}(D - HX) \tag{6}$$

$$X^a = X + \frac{A}{N-1}\frac{(HA)'}{P}(D - HX) \tag{7}$$

$$P = \frac{HA}{N-1} * (HA)' + R \tag{8}$$

The analysis formula Equation (7) can be rewritten in general form as:

$$X^a = X + K(D - HX) \tag{9}$$

where *K* is the Kalman gain expressed as:

$$K = \frac{A}{N-1}\frac{(HA)'}{P} \tag{10}$$

5. A TWIN EXPERIMENT

As outlined above in section 1 and 2, the hydrodynamics of the area of interest is highly complex. We therefore opt to test first the feasibility of the EnKF on a schematic 1D tidally driven estuary. This 1D case represents the basic characteristics of Malacca Strait, which connects Singapore waters with the Andaman Sea. An EnKF twin experiment is defined, in which EnKF adjusts (corrects) the disturbed M_2 tidal forcing back to the original undisturbed values or true solution. The schematized estuary is shown in Figure 1. This region has a width of 750 m, depth of 9 m and stretches from west to east. The eastern boundary, near Singapore Island is assumed to be a closed boundary, since major width is bounded by the land boundary. It is also assumed that at the open boundary the periodic water level is due to tidal component M2, since the M2 tidal component is dominant in this region. The time series of water level is at every 10 min time steps. Time series of M2 tidal predictions are assumed available at stations S1 - S4, as shown in Figure 1. Station S1 is chosen as an assimilation station and remaining stations S2, S3 and S4 are chosen as validation stations. True forcing at west boundary is:

$$\text{waterlevel (t)} = \mathrm{A} * \sin(2\pi\ \mathrm{t/T}) + \mathrm{w(t)} \tag{11}$$

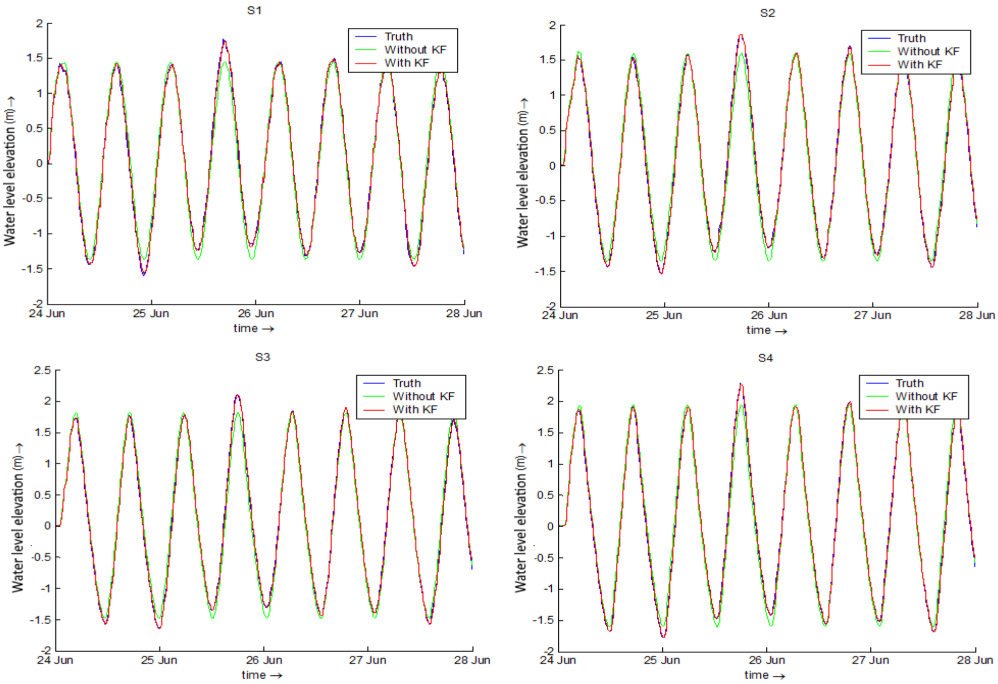


Figure 3. Water level at Station 1 (assimilation station), Station 2, 3, 4 (validation stations)

where A=1.5 m, T=12.42 h (M2) and w(t) is colored process noise which is represented as:

$$w(t) = \alpha\, w(t-1) + \eta(t-1) \qquad (12)$$

where

$$\alpha = \exp(-dt/C_t)\ ; C_t = 3\ h \qquad (13)$$

and η is white noise with

$$\eta \sim N(0,\sigma^2), \qquad (14)$$

$$\sigma = \sqrt{\left(\left(1-\alpha^2\right)\sigma_w^2\right)}\ ;\ \sigma_w=0.1\ m \qquad (15)$$

Tidal observation data at these stations are generated by perturbing the water level with colored noise and these are treated as true values. The time domain of water levels at each station with flow boundary conditions (M2) in Delft3D flow model are interpreted as the model run without data assimilation. The colored noise in the true values can be result of unknown physical process of the stochastic model and is estimated using the data assimilation tool (EnKF) combined with Delft3D model.

To simulate the nonlinear complex system, a portable data assimilation tool, OpenDA, is implemented which utilizes a set of interfaces that represent communication amongst models and associated perceptions. With the objective of improving the estimates of water level at the observation stations, the DA based EnKF is implemented.

6. RESULTS AND DISCUSSION OF THE TWIN EXPERIMENTS

In this section, we validate and discuss the results of the twin experiment applied to the schematized estuary described in Section 5 with 32 ensembles with standard deviation of measurement noise and process noise at 0.05 m and 0.05 m, respectively. Higher ensemble members resulted in slightly improved predictions, but at a cost of high computation time. Hence as a trade-off between the accuracy and the computation time, the present work is limited to 32 ensemble members. Observation location, S1 is chosen as the assimilation station and the remaining stations are chosen as validation stations. The results shown in Figure 3 are for each station with true forcing boundary conditions (with colored noise). It can be observed that the station (S1) which is closest to the open boundary resembles the impact of the disturbance best, while it gets less intense at the station (S4) which is farther from the open boundary.

Several sensitivity runs are conducted systematically to study the effect of parameters like the ensemble size, process or model noise and measurement noise on the improvement of the model prediction results compared against the true observations. These improvements are expressed in terms of the RMSE values at each station (location). The variation of ensemble member size varying from 8 to 100 is shown in Figure 4.

It can be observed that an ensemble size of more than 16 will result in lower RMSE but at the expense of computational cost. Since station S1 is chosen as the assimilation station, an ensemble size of 32 gives the optimal RMSE at the measured station and validation stations.

The influence of the variance of the noise in the process and measurement is studied by varying the standard deviations from 0.001 m to 0.1 m with ensemble size 32. Figure 5 shows the effect of different levels of noise in updating the model values while the standard deviation of the other noise is kept at 0.05 m.

A detailed study is also carried out to analyze the influence of the number of observation stations. In this scenario, we vary the observation stations to predict and forecast the water level at the remaining stations. The ensemble size and standard deviation of noise are set as 32 and 0.05 m for all these scenarios. The prediction results at each location for different observation stations are shown in Figure 6. It can be observed that, the best estimation result is obtained with S1 as the observation station as expected, since station S1 is near the open boundary. The ensemble filter is also able to provide good estimation at other stations.

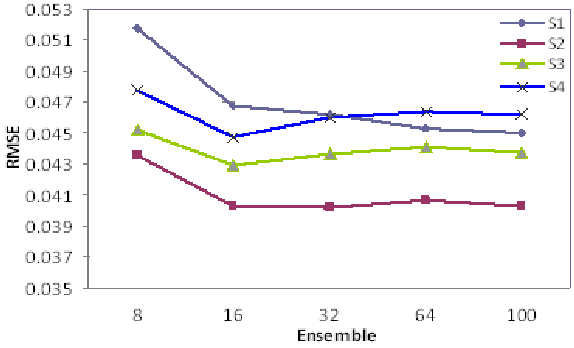
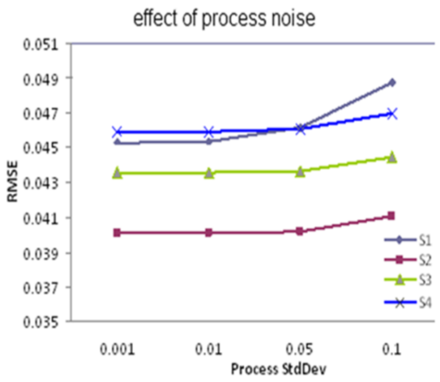


Figure 4. The RMSE (m) versus ensemble size at all the four stations



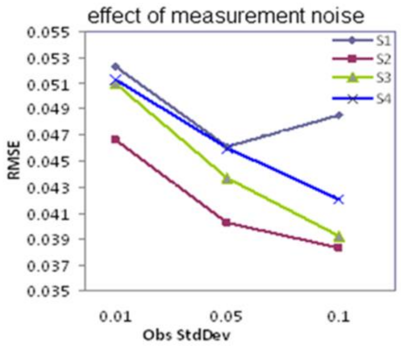


Figure 5. The RMSE (m) versus observed standard deviation at the four stations

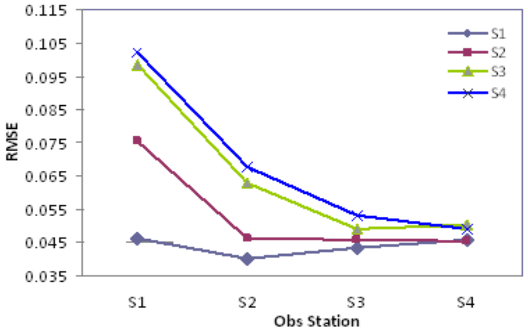


Figure 6. The RMSE (m) at each station with observation from different stations

7. CONCLUSIONS AND RECOMMENDATIONS

In this study, a systematic analysis has been undertaken to address the nonlinear interaction of large water bodies for accurately predicting the tides and surges in the Singapore regional waters. The stations are widely spread in the domain of study and variation of outcome clearly presents the sensitivity of domain. From the above analysis of a twin experiment on the enclosed estuary, it is observed that model predictions by data assimilation are very close to the observed (true) values. It was found that 32 ensembles provided optimal predicted water levels and an increase in the number of ensembles beyond 32 resulted in marginal enhancement but at the cost of higher computational effort. In this twin experiment, the influence of varying the process noise is negligible whereas measurement noise has great influence on the model predicted water levels. Overall, the above analysis show that the water levels predicted by OpenDA - EnKF are very much satisfactory and predicts values close to the true values. It is recommended that the technique be next applied to optimize the Andaman Sea tidal boundary forcing that is applied in the Singapore Region Model.

8. ACKNOWLEDGEMENT

The research presented in this work was carried out as part of Singapore-Delft Water Alliance (SDWA)'s “Must-Have Box” research program (R-264-001-003-272).

9. REFERENCES

1. Sabbagh-Yazdi, S.-R. and Zounemat-Kermani, M., "Numerical investigation of island effects on depth averaged fluctuating flow in the persian gulf", ***International Journal Of Engineering Transactions A Basics***, Vol. 20, No. 2, (2007), 117-124.
2. Afshin, H., Firoozabadi, B. and Rad, M., "Hydrodynamics analysis of density currents", ***International Journal of Engineering-Transactions B: Applications***, Vol. 21, No. 3, (2008), 211-220.
3. Bakhtiari, M., Veysi, S. and Ghassemi, H., "Numerical modeling of the stepped planing hull in calm water", ***International Journal of Engineering-Transactions B: Applications***, Vol. 29, No. 2, (2016), 236-240.
4. Kurniawan, A., Ooi, S.K., Hummel, S. and Gerritsen, H., "Sensitivity analysis of the tidal representation in singapore regional waters in a data assimilation environment", ***Ocean Dynamics***, Vol. 61, No. 8, (2011), 1121-1136.
5. Babovic, V. and Fuhrman, D.R., "Data assimilation of local model error forecasts in a deterministic model", ***International Journal for Numerical Methods in Fluids***, Vol. 39, No. 10, (2002), 887-918.
6. Karri. R.R.,, Rao, D. and Venkateswarlu, C., "Soft sensor based nonlinear control of a chaotic reactor", ***IFAC Proceedings Volumes***, Vol. 42, No. 19, (2009), 537-543.
7. Kalman, R.E., "A new approach to linear filtering and prediction problems", ***Journal of basic Engineering***, Vol. 82, No. 1, (1960), 35-45.
8. Kenarangui, R. and Mahdavi Tabatabayee, N., "On line electric power systems state estimation using kalman filtering (research note)", ***International Journal of Engineering***, Vol. 8, 4233-4240.
9. El Serafy, G.Y. and Mynett, A.E., "Improving the operational forecasting system of the stratified flow in osaka bay using an ensemble kalman filter–based steady state kalman filter", ***Water Resources Research***, Vol. 44, No. 6, (2008).
10. Verlaan, M., van Velzen, N., Hummel, S. and Gerritsen, H., "A generic toolbox for data-assimilation in numerical modeling", in 15th biennial conference of the joint numeric. (2010).
11. Badwe, A., Xuan, W., Karri, R.R., El, G. And Gerritsen, H., "Data assimilation for water levels and currents in the singapore region: An ensemble steady state kalman filtering approach".
12. El Serafy, G.Y., Gerritsen, H., Hummel, S., Weerts, A.H., Mynett, A.E. and Tanaka, M., "Application of data assimilation in portable operational forecasting systems—the datools assimilation environment", ***Ocean Dynamics***, Vol. 57, No. 4-5, (2007), 485-499.
13. Weerts, A.H., El Serafy, G.Y., Hummel, S., Dhondia, J. and Gerritsen, H., "Application of generic data assimilation tools (datools) for flood forecasting purposes", ***Computers & Geosciences***, Vol. 36, No. 4, (2010), 453-463.
14. Babovic, V., Karri, R., Wang, X., Ooi, S.K. and Badwe, A., "Efficient data assimilation for accurate forecasting of sea-level anomalies and residual currents using the singapore regional model", in Geophys Res Abstr. Vol. 13, (2011).
15. Karri, R.R., Wang, X. and Gerritsen, H., "Ensemble based prediction of water levels and residual currents in singapore regional waters for operational forecasting", ***Environmental Modelling & Software***, Vol. 54, (2014), 24-38.
16. Evensen, G., "Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics", ***Journal of Geophysical Research: Oceans***, Vol. 99, No. C5, (1994), 10143-10162.
17. Evensen, G., "The ensemble kalman filter: Theoretical formulation and practical implementation", ***Ocean Dynamics***, Vol. 53, No. 4, (2003), 343-367.
18. Houtekamer, P.L. and Mitchell, H.L., "Data assimilation using an ensemble kalman filter technique", ***Monthly Weather Review***, Vol. 126, No. 3, (1998), 796-811.
19. Karri, R.R., Badwe, A., Wang, X., El Serafy, G., Sumihar, J., Babovic, V. and Gerritsen, H., "Application of data assimilation for improving forecast of water levels and residual currents in singapore regional waters", ***Ocean Dynamics***, Vol. 63, No. 1, (2013), 43-61.
20. Xu, T. and Gomez-Hernandez, J.J., "Inverse sequential simulation: A new approach for the characterization of hydraulic conductivities demonstrated on a non-gaussian field", ***Water Resources Research***, Vol. 51, No. 4, (2015), 2227-2242.

Enhanced Predictions of Tides and Surges through Data Assimilation

TECHNICAL NOTE

R. R. Karri[a], V. Babovic[b]

[a] *Petroleum and Chemical Engineering, Universiti Teknologi Brunei, Brunei Darussalam*
[b] *Department of Civil and Environmental Engineering, National University of Singapore, Singapore*

PAPER INFO

Paper history:
Received 29 September 2016
Received in revised form 04 November 2016
Accepted 11 November 2016

Keywords:
Ensemble
Kalman filter
Data model integration
Data assimilation
OpenDA
Tides and surges

چکیده

آب منطقه‌ای در تنگه سنگاپور توسط پدیده‌های پیچیده هیدرودینامیکی در قالب نتیجه اثر ترکیبی سه حوزه بزرگ آبی یعنی دریای چین جنوبی، دریای آندامان و دریای جاوا شناسایی می‌شود. این امر منجر به ناهنجاری هایی در سطح آب و ایجاد جریان رسوبی می‌شود. مدل های هیدرودینامیکی عددی به طور کلی برای پیش بینی سطح آب در اقیانوس ها و دریاها استفاده می شوند. اما صحت آنها معمولا توسط عوامل متعدد، یعنی پیچیدگی های مرتبط با هندسه ساحلی و عدم قطعیت در جریان عواملی مانند (باد، فشار و جزر و مد اقیانوس عمیق) محدود شده است. مدل سازی دینامیک اقیانوس در تنگه مالاکا و آب های منطقه ای سنگاپور با توجه به حضور تعداد زیادی از جزایر کوچک و فعل و انفعالات جزر و مدیِ به شدت غیر خطی، مساله‌ای چالشی است. این پیچیدگی با توجه به عمق سنجی محلی و تغییرات هندسی در اطراف جزیره سنگاپور و اثرات هواشناسی در مقیاس های مختلف افزایش یافته است. این مطالعه به پیش بینی بهتر جزر و مد و امواج از طریق همانندسازی اطلاعات می‌پردازد. از طریق رابط کاربری قابل حمل OpenDA، فیلتر کالمن با یک مدل هیدرودینامیکی یکپارچه به منظور بهبود پیش بینی های مدل تلفیق می‌شود. برای ارزیابی حساسیت و بهبود مدل، یک آزمایش دوتایی طراحی شده است تا اثر مرزی جزر و مد را در خور نیمه محصور بهبود دهد. نتایج اصلی این مطالعه نشان می دهد که نتایج مدل به طور قابل توجهی در این رژیم جریان پیچیده بهبود یافته است.

***doi**: 10.5829/idosi.ije.2017.30.01a.04*