INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2009; **60**:651–667
Published online 10 September 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/fld.1910

Applying local model approach for tidal prediction in a deterministic model

Yabin Sun[1], Piyamarn Sisomphon[2,*,†], Vladan Babovic[1,2] and Eng Soon Chan[1,3]

[1]*Department of Civil Engineering, National University of Singapore, Singapore*
[2]*Singapore-Delft Water Alliance, National University of Singapore, Singapore*
[3]*Tropical Marine Science Institute, National University of Singapore, Singapore*

SUMMARY

In recent years, a practice of tidal prediction based on a deterministic model or by a time series forecasting model has been established. A deterministic model can predict tidal movement and capture the dynamics of the flow pattern over the entire domain. However, due to the simplification of model settings and near shore effects, the accuracy of the numerical model can diminish. Time series forecasting is capable of capturing the underlying mechanism that may not be revealed in the deterministic model simulation. However, such data-driven forecast fails to maintain accuracy with the progress of forecast horizon. In this paper, a scheme that combines the advantages of these two methods is introduced. The model errors are forecasted to different time horizons using a data-driven approach, and are then superimposed on the simulation results in order to correct the model output. Based on the proposed method, it is found that the accuracy is significantly improved with more than 50% of the errors removed on the average. Copyright © 2008 John Wiley & Sons, Ltd.

Received 20 February 2008; Revised 24 May 2008; Accepted 29 July 2008

KEY WORDS: local model; chaos theory; genetic algorithm; tidal prediction; error correction; deterministic model

1. INTRODUCTION

Tidal prediction is of prime importance for ship navigation and scheduling of harbor operations. Prediction could be made for time horizons from several hours to several days, which is helpful in the finalization of operational schedules and many other coastal activities. The current practice of tidal prediction is undertaken by either using a tidal predictive deterministic model or by a time series forecasting model. Each of the above two approaches has its own capabilities and deficiencies

*Correspondence to: Piyamarn Sisomphon, Strategic Research and Development, Deltares, The Netherlands.
†E-mail: ann.sisomphon@deltares.nl

Contract/grant sponsor: Singapore-Delft Water Alliance (SDWA); contract/grant number: R-264-001-001-272

of prediction. As long as the deterministic equations underlying the physical phenomena are known, in principle, they can be solved in a numerical model, such that the physics of the tidal movement can be predicted with a high prognostic capability [1, 2]. However, the accuracy of deterministic models depends on how accurate the model inputs are, especially regarding initial conditions and boundary conditions that govern the dynamics of the flow. Inaccurate model inputs will produce questionable model outputs irrespective of perfect knowledge of the governing laws. Alternatively, time series forecasting has been adopted in the field of tidal prediction. It ranges from linear methods, such as the autoregressive moving average approach, to more complex methods such as neural networks [3–5], radial basis functions [6], support vector machines [7] and genetic programming [3]. Another approach gaining popularity in the time series forecasting is local model (LM) inspired by chaos theory, which is capable of providing accurate short-term prediction [8, 9]. The Chaos theory reveals the underlying dynamics within the time series, but because it is extremely sensitive to the initial condition, the accuracy of prediction fades out when prediction horizon progresses.

In this paper, an attempt is made to optimize the advantages of the two approaches. The model errors between the model output and the observations are forecasted to the desired time horizons based on the revealed mechanisms in LM, which are then superimposed to the modeling results at individual tidal stations. By imposing the underlying dynamics of a chaotic time series on a deterministic predictive model, not only the accuracy but also the prognostic capability is improved. The efficiency of this combined scheme has been tested on the Singapore Regional Model (SRM) constructed within Delft3D modeling system. The error prediction is found to produce tidal forecast for longer prediction horizons with a higher accuracy than the direct forecasting from the numerical model.

2. DETERMINISTIC MODEL

Simulations are carried out with the Delft3D Modeling System developed by WL | Delft Hydraulics, the Netherlands. In particular, the hydrodynamic module FLOW is applied in this study. Delft3D-FLOW solves the Navier–Stokes equations for an incompressible fluid, under the shallow water and the Boussinesq assumptions. The numerical method of Delft3D-FLOW is based on finite differences, which solves the unsteady flow processes driven by tidal and meteorological forces on a curvilinear grid in multi-dimension (2D/3D). In this study the two-dimensional depth-integrated model is employed, and the basic equations can be described as given below [10].

The two-dimensional depth-integrated continuity equation is given by:

$$\frac{\partial h}{\partial t}+\frac{\partial}{\partial x}(h\overline{u})+\frac{\partial}{\partial y}(h\overline{v})=0 \tag{1}$$

where h is the local water depth, $\overline{u}$ and $\overline{v}$ indicate the depth-integrated fluid velocities in the x and y directions, respectively and t is the time. The two-dimensional momentum balance equation for fluid can be expressed as:

In the x direction

$$\frac{\partial}{\partial t}(h\overline{u})+\frac{\partial}{\partial x}(h\overline{u}\overline{u})+\frac{\partial}{\partial y}(h\overline{u}\overline{v})+\mathrm{g}h\frac{\partial}{\partial x}(h+z_b)+k_x h\left[\frac{\partial^2\overline{u}}{\partial x^2}+\frac{\partial^2\overline{v}}{\partial y^2}\right]-\frac{\tau_{bx}}{\rho}-\sum\frac{F_x}{\rho}=0 \tag{2}$$

DOI: 10.1002/fld

In the y direction

$$\frac{\partial}{\partial t}(h\overline{v})+\frac{\partial}{\partial x}(h\overline{vv})+\frac{\partial}{\partial y}(h\overline{vu})+gh\frac{\partial}{\partial y}(h+z_b)+k_y h\left[\frac{\partial^2\overline{v}}{\partial x^2}+\frac{\partial^2\overline{u}}{\partial y^2}\right]-\frac{\tau_{by}}{\rho}-\sum\frac{F_y}{\rho}=0 \tag{3}$$

in which g is the gravitational acceleration, z_b represents the bed level above reference datum, k_x and k_y are the effective dispersion coefficients representing the integration effects, ρ is the fluid density and F indicates the external driving forces e.g. wind, waves, Coriolis effect.

3. SINGAPORE REGIONAL MODEL

The dedicated SRM constructed within the Delft3D modeling system is employed in this study. This model has been developed with an intention to provide reliable hydrodynamic information about the water surrounding Singapore. As shown in Figure 1, the model domain with the highlighted area around Singapore Island, covers large parts of the seas around Singapore. Open boundaries are located in the Andaman Sea (north-west of Singapore), in the South China Sea (north-east of Singapore) and in the Java Sea (south-east of Singapore), where a total of eight tidal constituents (Q1, O1, P1, K1, N2 M2, S2 and K2) have been prescribed as boundary conditions. The bathymetry in the SRM is based on Admiralty charts. The maximum depth of the model is about 2000 m in the Andaman Sea and the maximum depth in the Singapore Straits is more than 160 m. The SRM grid consists of around 38 500 grid cells in the horizontal plane. Grid sizes vary from about 200 to 300 m around Singapore Island up to over 15 km at the open boundaries.

Singapore is located between two large water bodies: the South China Sea on the east and the Andaman Sea on the west. The water motion in the Singapore Straits is driven by tides coming from both sides, by mean sea level differences between seas and by the wind. Therefore, the hydrodynamics of Singapore water is complex. As presented in Figure 2, four tidal stations are studied in this paper, which are located at Pasir Panjang (1.30 N, 103.76 E), Raffles (1.16 N, 103.74 E), Sembawang (1.47 N, 1.84 E) and Horsburgh (1.33 N, 104.41 E). The observations at these stations are obtained from the 1999 Singapore Tide Tables and Port Information [11].

The simulation covers one-year period starting from 1 January 0:00 1999 to 31 December 23:00 1999, producing 8760 data for each grid point. After neglecting the first 480 time steps to avoid initialization effect, the SRM output is compared with the observations at these four stations, and the results are shown in Table I in terms of mean absolute error (MAE), root mean square error (RMSE) and correlation coefficient (r) defined as follows:

$$\mathrm{MAE}=\frac{\sum|x_i-x_i'|}{N} \tag{4}$$

$$\mathrm{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-x_i')^2} \tag{5}$$

$$r=\frac{\sum(x_i-\overline{x}_i)(x_i'-\overline{x}_i')}{\sqrt{\sum(x_i-\overline{x}_i)^2\sum(x_i'-\overline{x}_i')^2}} \tag{6}$$

$$\overline{x}_i=\frac{\sum x_i}{N},\quad \overline{x}_i'=\frac{\sum x_i'}{N} \tag{7}$$

DOI: 10.1002/fld

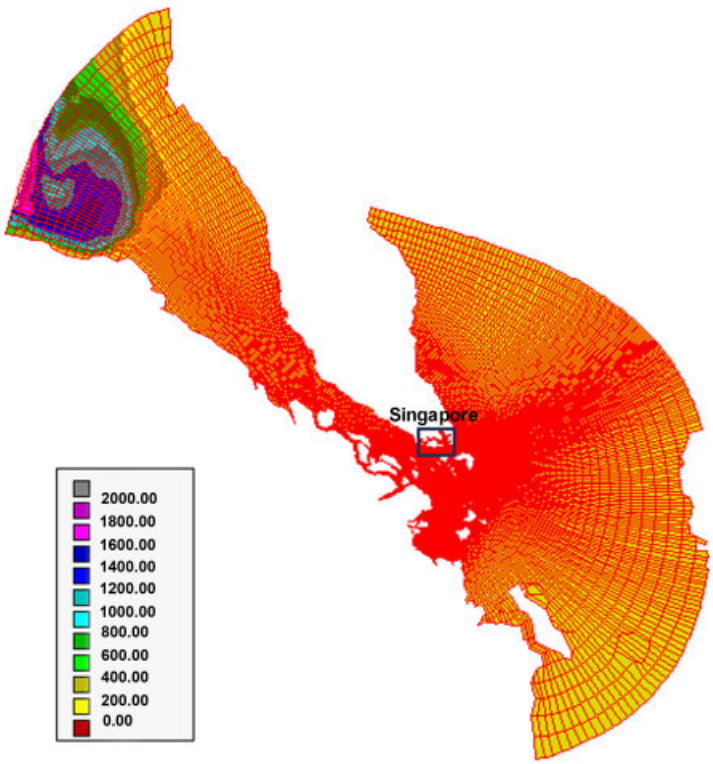

Figure 1. Extent and bathymetry of Singapore Regional Model.

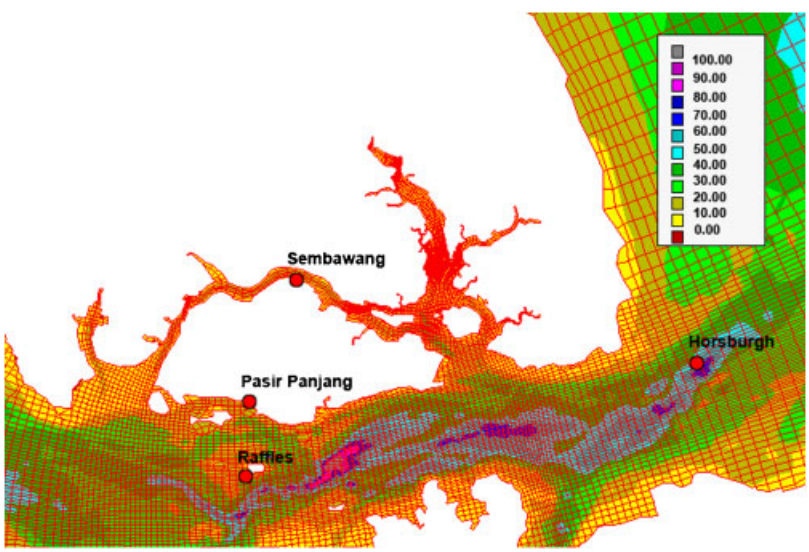

Figure 2. The tidal stations in the Singapore Regional Model (SRM).

DOI: 10.1002/fld

Table I. Statistics of the model errors at the tidal stations.

Tidal station	Pasir Panjang	Raffles	Sembawang	Horsburgh
MAE (cm)	15.32	13.90	14.07	8.90
RMSE (cm)	18.17	16.66	17.40	10.91
r	0.92	0.91	0.89	0.95

where N is the length of the time series, x_i represents the observed value and x_i' is the SRM output.

In SRM, only eight primary constituents are specified at the open boundary as the external force, which may be inadequate to describe the complexity of the dynamics in Singapore water. Among all these four tidal stations, Horsburgh is located in the open sea, where the water depth is about 60 m and the water dynamics can be better described by the eight tidal constituents than the other three stations. Compared with Horsburgh, Pasir Panjang, Raffles and Sembawang are closer to the coast with a water depth ranging from 10 to 20 m, where advection, large amplitude to depth ratio and bottom friction give rise to non-linear interactions. Therefore, the SRM output is more accurate at Horsburgh despite the coarser grid in that area of computational domain. In addition, due to the shoaling effect, the model errors at Pasir Panjang, Raffles and Sembawang have more nonlinearities than the errors at Horsburgh. Examples of model outputs at the tidal stations are plotted in Figure 3, accompanied by observations and model errors, from which discrepancies between model outputs and observations can be noticed especially at the tidal level extrema.

4. CHAOS THEORY

Recent developments in non-linear dynamics have demonstrated that irregular or random behavior in natural systems may arise from purely deterministic dynamics with unstable trajectories. Although some observations might appear random, beneath their random behavior may lie an order or pattern. Such type of non-linear dynamical systems, which are highly sensitive to initial conditions, are popularly known as chaotic systems. According to Williams [12], chaos is a sustained and disorderly looking evolution that satisfies certain special mathematical criteria and occurs in a deterministic non-linear system.

4.1. Embedding theorem

Takens time-delay embedding theorem [13] paved the way for the analysis of chaotic time series in chaotic systems. The theorem essentially states that the underlying structure of a complex, multi-dimensional system can be viewed using a projection from a single variable in the phase space, which is an embedded space with dimensions consisting of various time lags of the variable itself.

The embedding theorem establishes that, given a scalar time series x_i from a dynamical system, it is possible to reconstruct a phase space in terms of the phase space vector $\mathbf{x}_i$ expressed as

$$\mathbf{x}_i = (x_i, x_{i-\tau}, \ldots, x_{i-(m-1)\tau}) \tag{8}$$

where m is the embedding dimension and τ is the time delay. According to the embedding theorem, the underlying structure cannot be seen in the space of the original scalar time series, rather only when unfolded into a phase space or embedded space.

DOI: 10.1002/fld

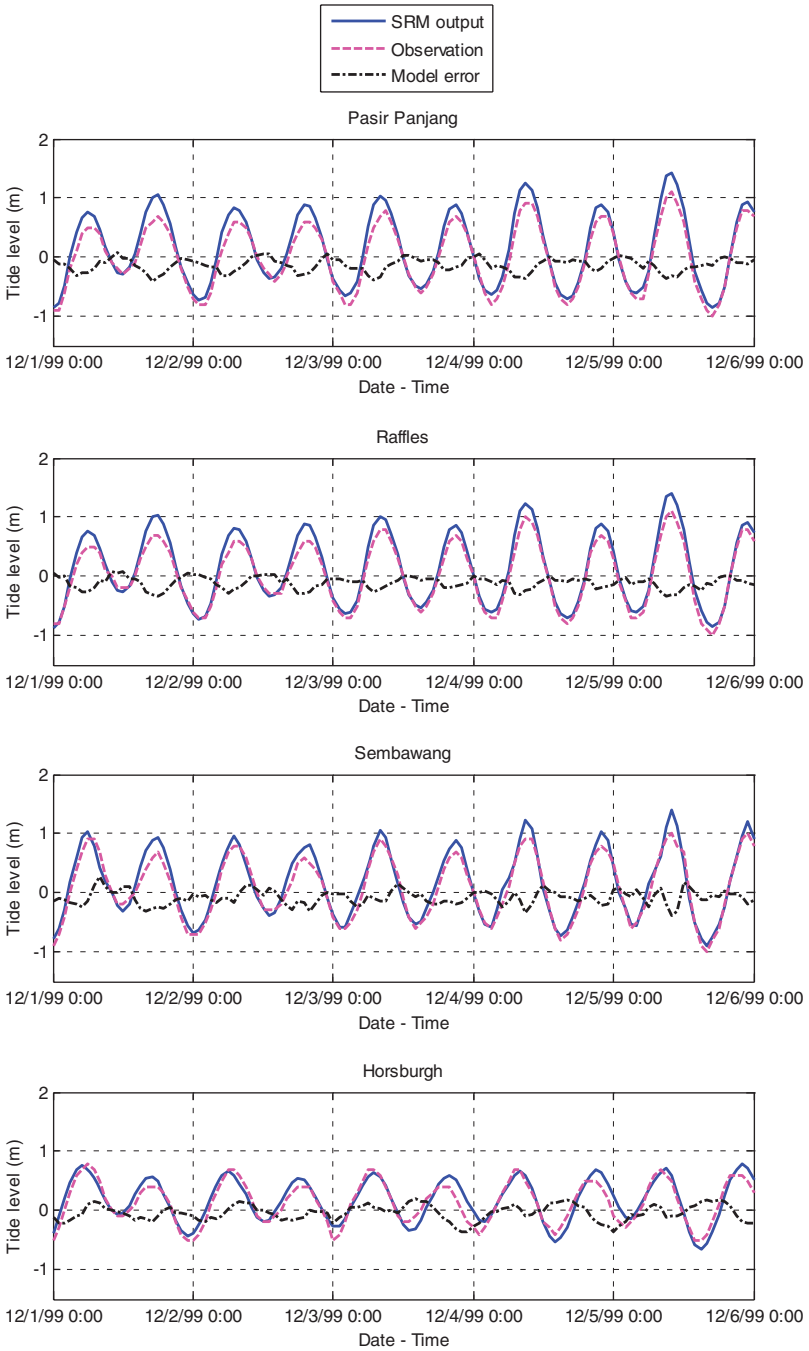

Figure 3. SRM output, observation and model error.

DOI: 10.1002/fld

In the phase space prediction model, the basic idea is to set a functional relationship between the current state $\mathbf{x}_t$ and the future state $\mathbf{x}_{t+T}$ in the form

$$\mathbf{x}_{t+T} = f_T(\mathbf{x}_t) \tag{9}$$

where T is referred to as lead time or prediction horizon, and the problem now is limited to find a good expression for the mapping function f_T.

4.2. Local model approach

LM [14, 15] is an effective method of simulating the evolution of a dynamical system by means of local approximation, using only the most similar trajectories from the past to make predictions for the future [9, 16–19]. A conceptual sketch of this algorithm is depicted in Figure 4, followed by the steps description in LM.

Step 1: Embedding the time series into a phase space.

Step 2: Finding k nearest neighbors in the phase space.

To predict a future state $\mathbf{x}_{t+T}$, a Euclidean metric is imposed on the phase space to find the k nearest neighbors of the current state $\mathbf{x}_t$, which are denoted by $\mathbf{x}_n$ $(n = 1, 2, \ldots, k)$.

Step 3: Calculating the 'expected' future state.

Having constructed the phase space and pooled the k nearest neighbors of the current state $\mathbf{x}_t$, the 'expected' vector of the future state $\mathbf{x}_{t+T}$, denoted as $\hat{\mathbf{x}}_{t+T}$, can be estimated through averaging as [14, 15, 20]

$$\hat{\mathbf{x}}_{t+T} = \left(\sum_{n=1}^{k} \mathbf{x}_{n+T} \right) \Big/ k \tag{10}$$

Step 4: Deriving the forecast scalar value.

In the phase space, the 'expected' future state $\hat{\mathbf{x}}_{t+T}$ can be expressed in the form of Equation (8) as

$$\hat{\mathbf{x}}_{t+T} = (\hat{x}_{t+T}, \hat{x}_{t+T-\tau}, \ldots, \hat{x}_{t+T-(m-1)\tau}) \tag{11}$$

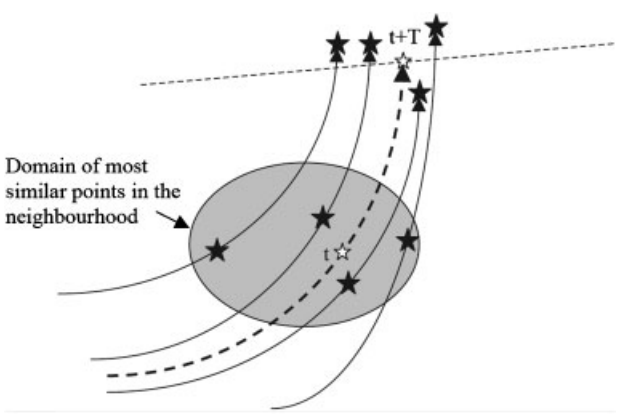

Figure 4. Conceptual sketch of forecasting using nearest neighbors. The black stars are the nearest neighbors to the white star representing the predicted value (adapted from Sannasiraj *et al.* [9]).

DOI: 10.1002/fld

and the predicted scalar values $\hat{x}_{t+T}, \hat{x}_{t+T-\tau}, \ldots$ in the time series x_i can be retrieved according to the structure.

Although LM makes use of a linear approximation for each separate prediction, the resulting overall model can be highly non-linear, as each of these linear approximations is made for each separate neighborhood.

When making a LM forecast, the first step is to unfold the time series into a phase space, which typically involves the selection of an embedding dimension m and a time delay τ. In the traditional standard approach, false nearest neighbors and average mutual information analyses are recommended to determine these parameters [21]. However, the standard approach has shown to provide suboptimal choices of embedding parameters, although it has more meaning from the theoretical point of view. Therefore, in this paper, an alternate inverse approach based on genetic algorithm (GA) is employed, which has demonstrated significant improvements over the standard approach (see, e.g. References [18, 19, 22–24]).

4.3. Genetic algorithm

GA [25] is a parameter search procedure based on the mechanics of natural genetics, which combines the Darwinian theory of evolution with a random, yet structured information exchange among a population of artificial chromosomes.

In the present paper, GA is used for simultaneously optimizing embedding dimension m, time delay τ and the number of nearest neighbors k. The flow diagram for GA is presented in Figure 5,

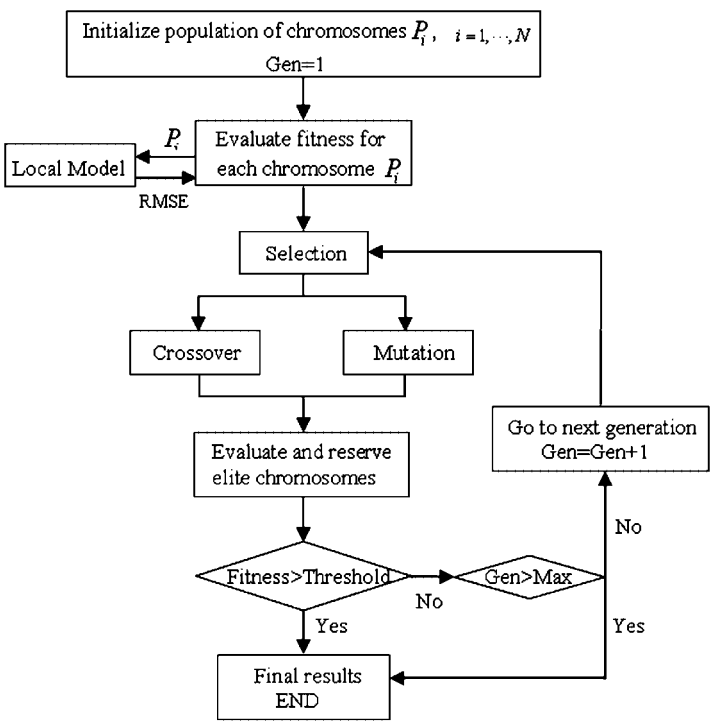

Figure 5. Flow diagram for genetic algorithm.

DOI: 10.1002/fld

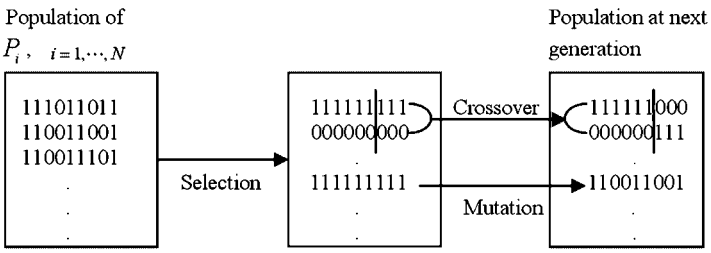

Figure 6. Schematic illustration of the evolving process in genetic algorithm.

while Figure 6 illustrates the evolving process in GA. In principle, an initial population of chromosomes $P_i = \{m, \tau, k\}$, where m, τ and k are represented by binary bits, is randomly generated within the specified ranges of parameters, and is allowed to evolve through the following process:

- *Selection*: A simple scheme is employed to select the chromosomes to reproduce offspring according to their respective fitness. The chromosome with a higher fitness has a better chance of being selected. For every chromosome, LM is executed to evaluate the fitness in terms of RMSE.
- *Crossover*: Some portion of a pair of chromosomes selected from the population is exchanged according to some constraints in order to generate two new sets of parameters.
- *Mutation*: One individual chromosome selected from the population is transformed to a new individual by inverting some of its binary values.

The process is continued until an entirely new population is generated with the hope that the fitter parents will create a better generation of children, such that the average fitness of the population will tend to increase with each new generation. The fitness of each child in the new generation is evaluated, and the process of selection, crossover and mutation is repeated. Successive generations are created until the user-defined threshold for the fitness or the number of maximum generation is reached.

5. RESULTS AND DISCUSSION

The present study aims at predicting the model errors to eight prediction horizons ranging from 2 to 96 h using LM. Among the total 8760 data in 1999, the first 480 steps (1 January 0:00–20 January 23:00) are discarded due to the initialization effect, and time steps from 481 to 7320 (21 January 0:00–1 November 23:00) are used to train the LM in determining the optimal m, τ and k for each prediction horizon required at each station, while the remaining data from 7321 to 8760 (2 November 0:00–31 December 23:00) are used as validation data to testify the efficiency of LM.

Following the idea of the embedding theorem, GA has been applied for the optimization of LM parameters. The specified ranges of genetic algorithm parameters are presented in Table II. It has been verified that GA already converges based on these parameters, and a further increase in the search scope will not significantly increase the efficiency of prediction. The optimized embedding parameters m, τ and k are presented in Table III, which are unique for each tidal station and different for various prediction horizons as well.

DOI: 10.1002/fld

Table II. Parameter settings in the genetic algorithm.

Parameter	Embedding dimension (m)	Time lag τ	No. of nearest neighbors k	Population size	Maximum generation
Range	1–10	1–50	1–100	10	200

Table III. Embedding parameters (m, τ, k) in local model.

Embedding parameters, I (h)	Pasir panjang			Raffles			Sembawang			Horsburgh		
	m	τ	k	m	τ	k	m	τ	k	m	τ	k
$T=2$	2	1	30	2	1	36	2	1	28	3	1	23
$T=6$	4	19	10	5	19	14	4	3	50	5	19	37
$T=12$	5	7	28	6	13	28	5	13	29	3	13	25
$T=24$	4	4	36	4	15	59	3	12	48	6	21	37
$T=36$	6	13	47	7	17	33	4	15	33	8	12	40
$T=48$	6	22	25	5	11	76	7	21	23	7	31	34
$T=72$	4	36	49	5	44	41	5	44	46	9	44	28
$T=96$	5	30	36	8	36	34	6	34	39	6	43	53

Figure 7 illustrates examples of error forecasting using LM with prediction horizon fixed to 2 h. The 2-h forecast agrees well with the model errors, producing trivial residual errors oscillating around zero. The corresponding reconstructed phase spaces are plotted in Figure 8, which are visualizable as embedding dimension $m \leqslant 3$. Clear patterns are revealed in the reconstructed phase space out of the original chaotic time series, based on which the model errors are correspondingly predicted. The underlying structure of the state vectors can be best viewed in a two-dimensional phase space for Pasir Panjang, Raffles and Sembawang, while the optimal embedding dimension is three for Horsburgh. As shown in Figure 8, the state vector points in the phase space of Horsburgh are a little denser than the vectors in the other three phase spaces, which makes the trajectory less traceable and the model errors less predictable. Depending on the patterns revealed in the phase spaces, more than 70% of errors have been removed from the SRM output for Pasir Panjang, Raffles and Sembawang, while more than 60% for Horsburgh, when $T=2$ h.

The prediction horizon is extended up to 96 h, which is normally sufficient for emergency offshore operations. Examples of error forecast at Pasir Panjang are depicted in Figure 9. As anticipated, the accuracy of prediction decreases when the prediction horizon increases. However, even when $T=96$ h, the LM forecast successfully resolves rising and falling tendencies of the model errors, and the residual error remains fluctuating about zero, which removes almost 50% of the errors. Moreover, the correlation coefficient between the corrected model output and the observations maintains larger than 0.97 for all the prediction horizons compared with the original 0.92 without error correction. This is further confirmed by the corresponding scatter diagrams of the forecasted tidal and the observations depicted in Figure 10. The direct tidal forecast from SRM is found to over-predict the reality, while the error correction scheme is capable of reducing the scatter and sustaining good accuracy over longer prediction horizons.

Table IV lists the error forecasting efficiency of LM through evaluating the respective residual MAE, RMSE, as well as the correlation coefficient after error correction. The evolution of residual

DOI: 10.1002/fld

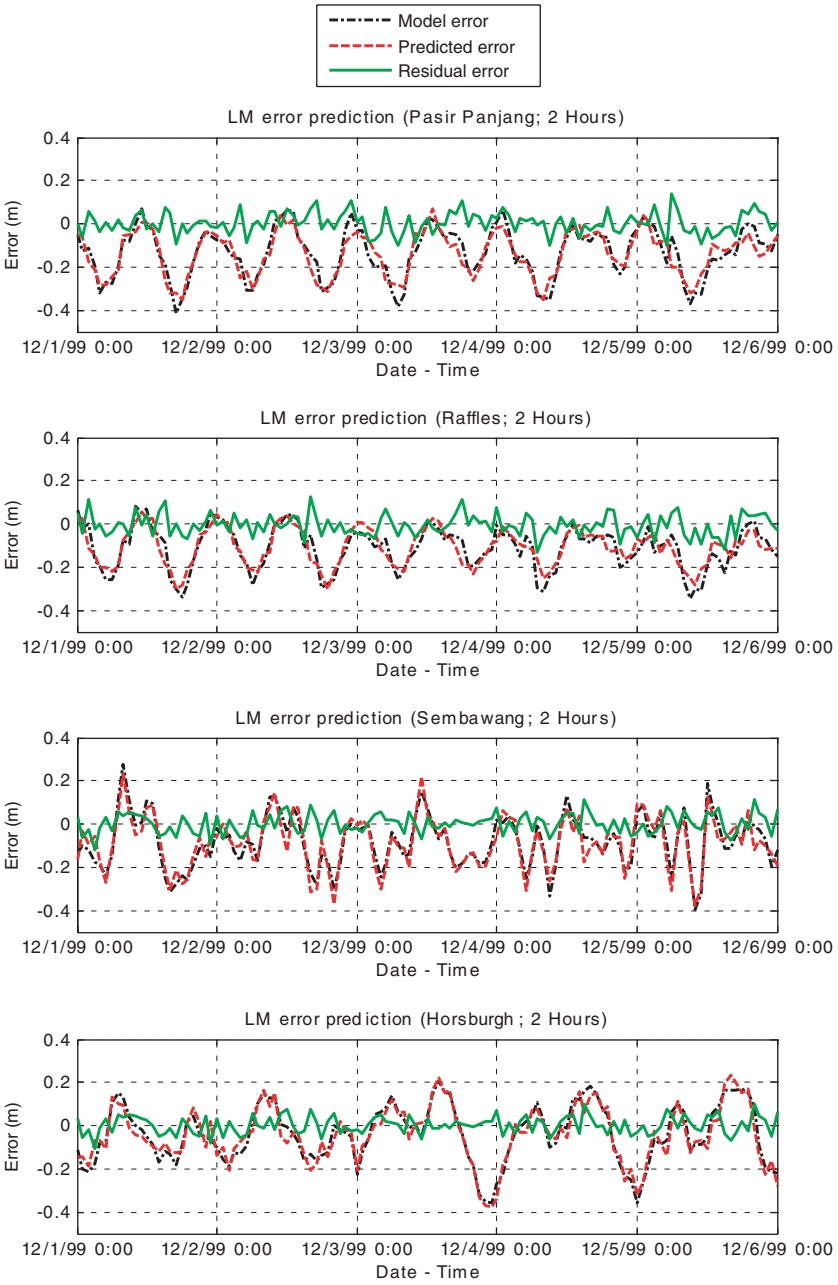

Figure 7. Example of error forecast using local model when $T = 2$h.

DOI: 10.1002/fld

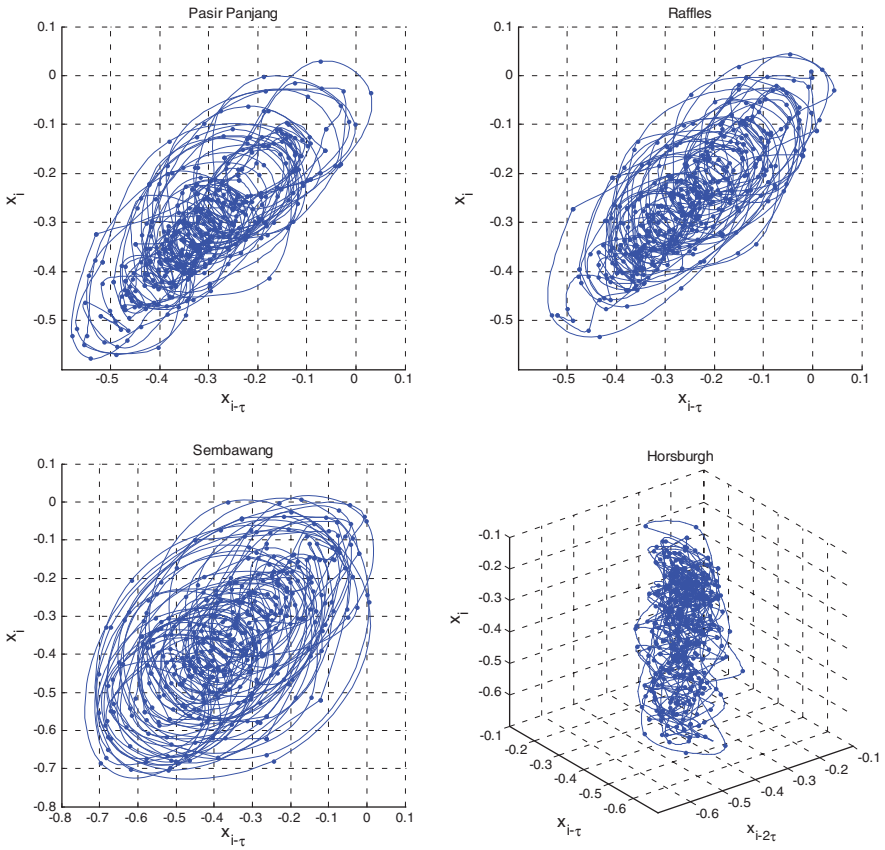

Figure 8. Reconstructed phase space when $T = 2$ h.

RMSE against forecast horizons in the corrected model at the tidal stations is plotted in Figure 11, with the RMSE before correction. It is observed that errors are reduced to a great measure, and the curves have a slight increasing trend, which is reasonable due to the decreasing predictability of model errors with the increasing prediction horizons. As stated earlier, the model errors at the stations located close to the coast are more nonlinear than the ones at the stations in the open sea. Therefore, LM performs relatively better at Pasir Panjang, Raffles and Sembawang than at Horsburgh.

6. CONCLUSIONS

The present paper describes an efficient error correction scheme combining a deterministic numerical model Delft3D and a time series prediction algorithm LM. Based on the utilization of Delft3D and LM, the combined scheme is capable of providing highly accurate forecast with a longer prediction horizon, and has been successfully applied in improving the tidal prediction accuracy of SRM.

DOI: 10.1002/fld

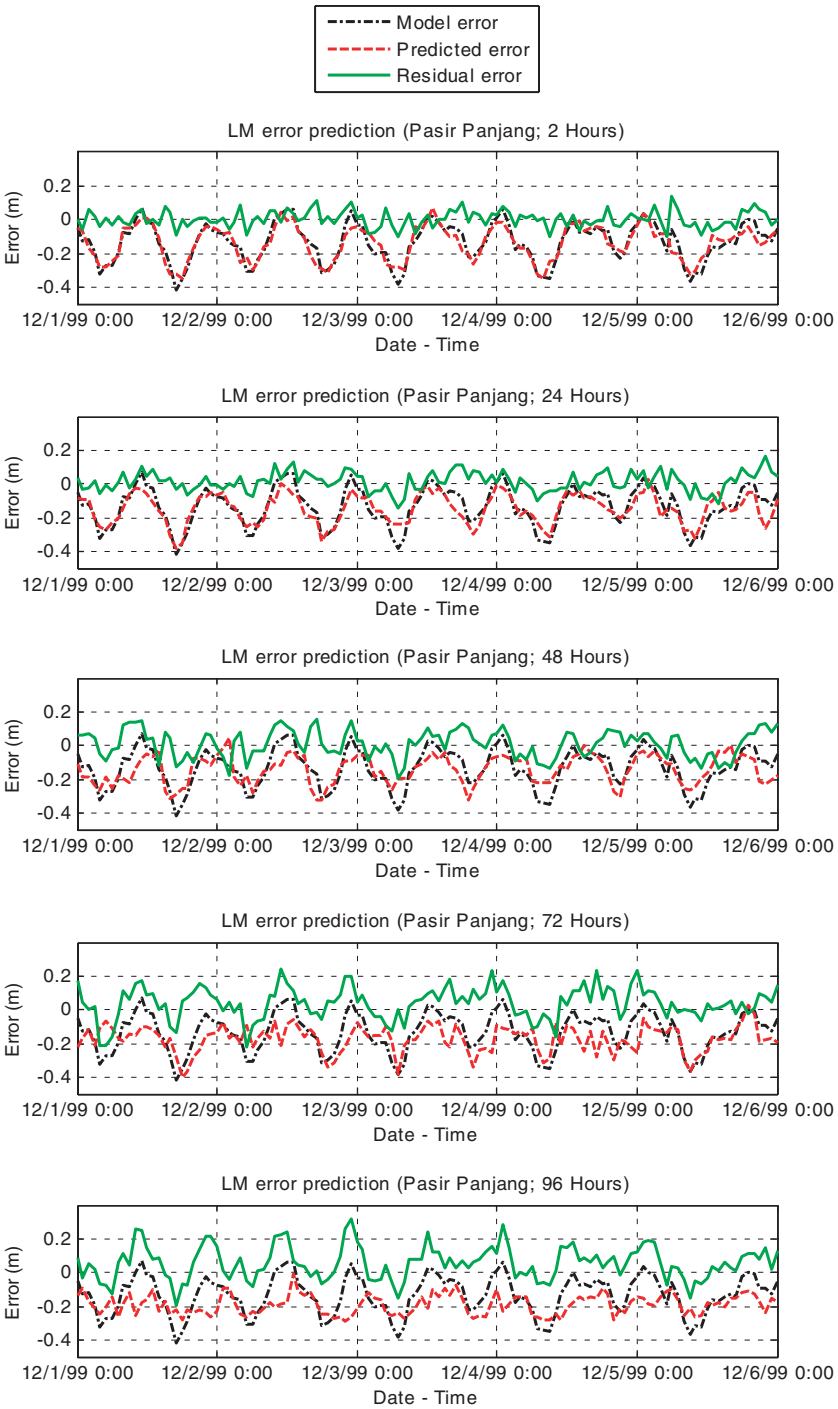

Figure 9. Example of error forecast using local model at Pasir Panjang.

DOI: 10.1002/fld

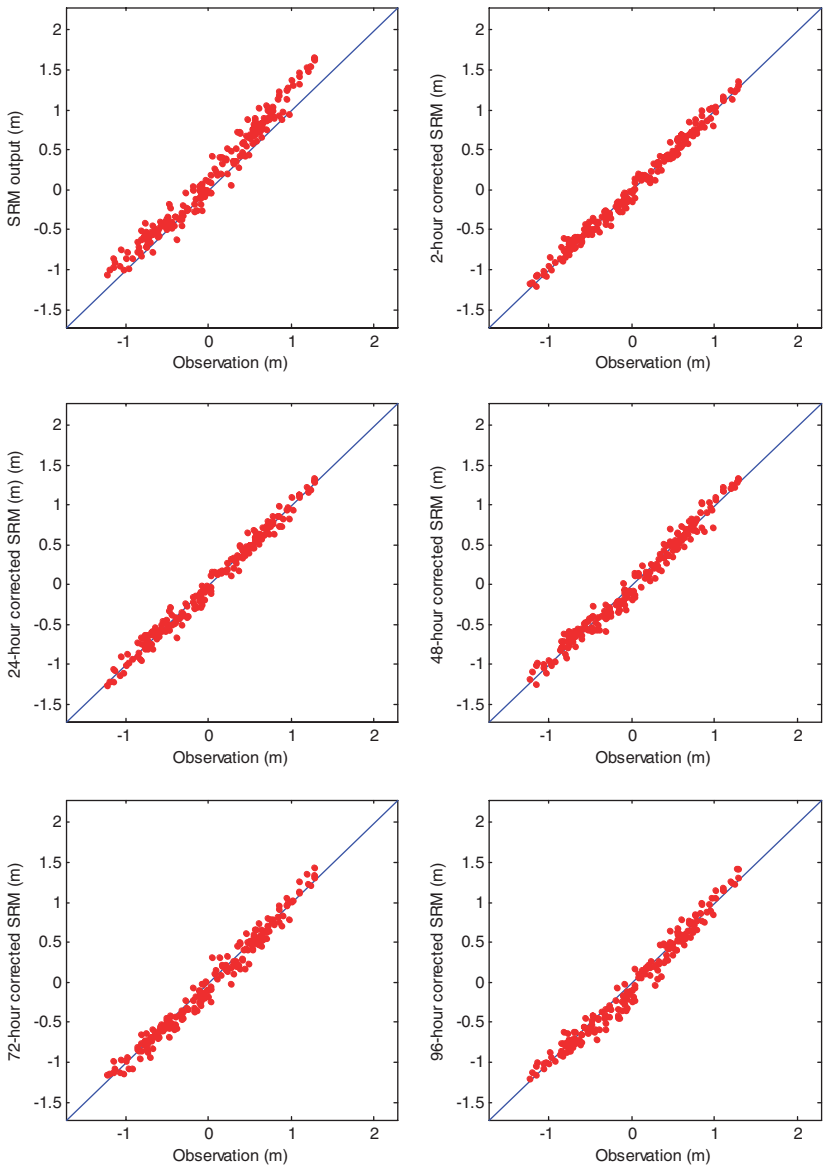

Figure 10. Scatter diagrams of tidal forecast and the observation at Pasir Panjang.

Compared with the observed tidal levels, SRM tends to over-predict the physical tidal characteristics due to the simplified model settings and near shore effects. The nonlinear time series forecasting model, LM, then is imposed on the numerical predictive results as an error prediction algorithm. The optimal embedding parameters in LM are determined by GAs, and model errors are predicted to different prediction horizons based on the structure revealed in the embedded phase space. The resultant corrected model output is shown to be more accurate in terms of smaller

DOI: 10.1002/fld

Table IV. Summary of residual errors and correlation coefficient at the tidal stations.

	Pasir Panjang			Raffles			Sembawang			Horsburgh		
Station	MAE	RMSE		MAE	RMSE		MAE	RMSE		MAE	RMSE	
T (h)	(cm)	(cm)	r	(cm)	(cm)	r	(cm)	(cm)	r	(cm)	(cm)	r
Model	15.32	18.17	0.92	13.90	16.66	0.91	14.07	17.40	0.89	8.90	10.91	0.95
2	3.75	4.74	0.99	3.87	4.88	0.99	3.80	4.78	0.99	3.40	4.28	0.99
6	3.98	5.03	0.99	4.12	5.23	0.99	4.12	5.19	0.99	3.88	4.89	0.99
12	4.33	5.41	0.99	4.44	5.60	0.99	4.69	5.90	0.99	4.20	5.29	0.99
24	4.57	5.79	0.99	4.63	5.82	0.99	5.25	6.53	0.98	4.22	5.76	0.99
36	5.66	7.16	0.99	5.57	7.08	0.99	5.70	7.08	0.98	4.42	5.80	0.99
48	6.11	7.86	0.98	5.80	7.43	0.98	5.70	7.20	0.98	4.70	6.26	0.99
72	7.17	9.34	0.98	6.67	8.35	0.97	6.52	8.19	0.97	5.05	6.70	0.98
96	7.69	9.70	0.97	7.33	9.41	0.97	7.79	9.80	0.96	5.92	7.99	0.98

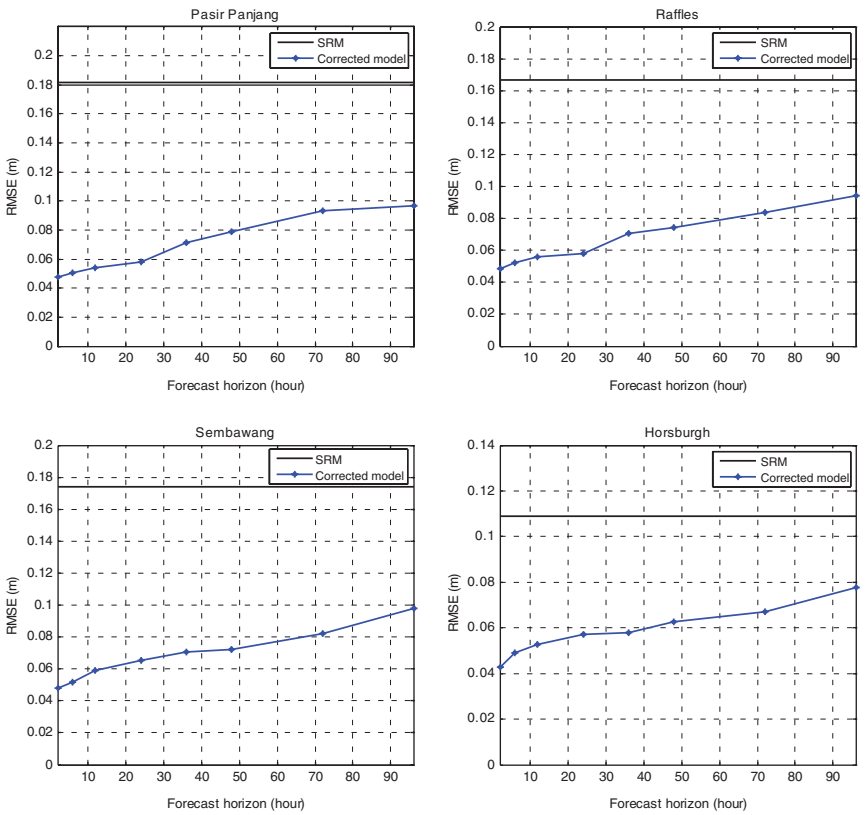

Figure 11. Comparison of RMSEs versus the prediction horizons.

DOI: 10.1002/fld

errors and higher correlation. At Pasir Panjang, more than 70% of the RMSE has been removed from the SRM, when $T = 2$ h and the error reduction is still almost 50% for a 96-h forecast. The correlation coefficient also increased from 0.92 to 0.99 for a 2-h forecast and 0.97 for a 96-h forecast. Additionally, the error correction scheme reduces the scatter and maintains good accuracy up to 96 h. When the prediction horizon progresses, it becomes more intractable to capture the trajectories of the state vectors in the unfolded phase space. This makes the model error time series less predictable. Therefore, the performance of LM deteriorates slightly with increasing prediction horizons.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support and the contributions of this project to the Singapore-Delft Water Alliance (SDWA). The research presented in this work was carried out as part of the Singapore-Delft Water Alliance's (SDWA) research programme (R-264-001-001-272).

REFERENCES

1. Pugh DT. *Tides*, *Surges and Mean Sea-Level*. Wiley: Chichester, 1996.
2. Palacio C, Winter C, Mayerle R. Set-up of a hydrodynamic model for the meldorf bight. *Proceedings of World Water and Environmental Resources Congress*, Orlando, 2001.
3. Babovic V. *Emergence*, *Evolution*, *Intelligence*: *Hydroinformatics*. Balkema: Rotterdam, 1996.
4. Minns AW. Artificial neural networks on subsymbolic process descriptors. *Ph.D. Thesis*, Balerma, Rotterdam, 1998.
5. Babovic V, Canizares R, Jensen HR, Klinting A. Neural networks as routine for error updating of numerical models. *Journal of Hydraulic Engineering* 2001; **127**:181–193.
6. Kantz H, Schreiber T. *Nonlinear Time Series Analysis*. Cambridge University Press: Cambridge, 1997.
7. Cristianini N, Shawe-Taylor J. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press: Cambridge, 2000.
8. Babovic V, Keijzer M. Forecasting of river discharges in the presence of chaos and noise. In *Coping with Floods*: *Lessons Learned from Recent Experiences*, Marsalek J (ed.). NATO ARW Series. Kluwer Academic Publishers: Dordrecht, 1999.
9. Sannasiraj SA, Babovic V, Chan ES. Local model approximation in the real time wave forecasting. *Coastal Engineering* 2005; **52**:221–236.
10. WL|Delft Hydraulics. *Delft3D-FLOW User Manual*, version 3.12. Delft, 2005.
11. *Singapore Tide Tables and Port Information*. Hydrographic Department of Singapore. Singapore, 1999. ISSN: 0129-0282.
12. Williams GP. *Chaos Theory Tamed*. Taylor & Francis: London, 1997.
13. Takens F. Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence*, Rand DA, Yong L-S (eds). Lecture Notes in Mathematics, vol. 898. Springer: Berlin, 1980; 366–381.
14. Farmer JD, Sidorowich JJ. Predicting chaotic time series. *Physical Review Letters* 1987; **59**:845–848.
15. Casdagli M. Nonlinear prediction of chaotic time series. *Physica D* 1989; **35**:335–356.
16. Keijzer M, Babovic V. Error correction of a deterministic model in Venice lagoon by local linear models. *Proceedings of Modelli Complessi e Metodi Computazionali Intensivi per la Stima e la Previsione*, Venice, 1999.
17. Babovic V, Keijzer M, Bundzelm M. From global to local modeling: a case study in error correction of deterministic models. *Proceedings of the 4th International Conference on Hydroinformatics*, Iowa City, 2000.
18. Babovic V, Keijzer M, Stefannson M. Optimal embedding using evolutionary algorithms. *Proceedings of the 4th International Conference on Hydroinformatics*, Iowa City, 2000.
19. Fuhrman DR. Data assimilation and error prediction using local models. *M.Sc. Thesis*, International Institute for Infrastructural, Hydraulic and Environmental Engineering (IHE), Delft, 2001.
20. Sugihara G, May RM. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature* 1990; **344**:734–741.
21. Abarbanel HDI. *Analysis of Observed Chaotic Data*. Springer: Berlin, New York, 1998.

DOI: 10.1002/fld

22. Phoon KK, Islam MN, Liaw CY, Liong SY. A practical inverse approach for forecasting nonlinear hydrological time series. *Journal of Hydrologic Engineering* 2002; **7**:116–128.
23. Babovic V, Sannasiraj SA, Chan ES. Error correction of a predictive ocean wave model using local model approximation. *Journal of Marine Systems* 2005; **53**:1–17.
24. Liong SY, Doan CD. Derivation of effective and efficient data set for training forecasting model. *Proceedings of the 13th IAHR-APD Congress*, Singapore, 2002.
25. Holland JH. *Adaptation in Natural and Artificial Systems*. University of Michigan Press: Ann Arbor, MI, 1975.

DOI: 10.1002/fld