Urban flash flood forecast using support vector machine and numerical simulation

Authors
J. Yan[1], J. Jin[2], F. Chen[1], W. Wang[1], G. Yu[1], H. Yin[3].
[1]DHI China, Urban Water Department, Shanghai, China.
[2]ZheJiang College Of Construction, Department of urban construction, Hangzhou, China.
[3]Tongji University, College of Environmental Science and Engineering, Shanghai, China.

Keywords
Forecast
Numerical simulation
Support vector machine
Urban flood

Abstract Text

Summary

In order to provide urban flood early warning effectively, two support vector machine (SVM) models, using a numerical model as data producer, were developed to forecast the flood alert and the maximum flood depth respectively. An application in the urban area of Jinlong River Basin, Hangzhou, China, showed the superiority of the proposed models. Statistical results based on the comparison between the results from SVM models and numerical model, proved that the SVM models could provide accurate forecast for estimating urban flood. For all the rainfall events tested with an identical desktop, the SVM models only took 2.1 milliseconds while the numerical model took 25 hours.

Introduction

Urban flood is a serious problem in cities worldwide, especially with increasing climate change and urbanization. Infrastructure damages and property losses caused by extreme rain event could be alleviated with the aid of urban flood forecast system.

Numerical models based on physical process are commonly used as urban flood simulation tools (Bach et al. 2014). However, they usually require sizable computing time (Casulli 2009). In order to achieve urban flood forecast over short temporal scale, support vector machine (SVM) supported with available data could be an alternative solution. The present study was aimed to develop an urban flood forecast framework combined numerical model based on MIKE FLOOD (DHI 2016) with SVM model.

Methods and Materials

Study area and data set description

The study area is the urban area of Jinlong River Basin, Hangzhou, China, with an area of 4.5 km^2. In the study area, 3 rainfall gauge stations and 2 water level gauge stations were installed (Fig. 1). The monitoring data during the year of 2013, the terrain data, Jinlong river data and local rainfall statistic data were employed in this study.

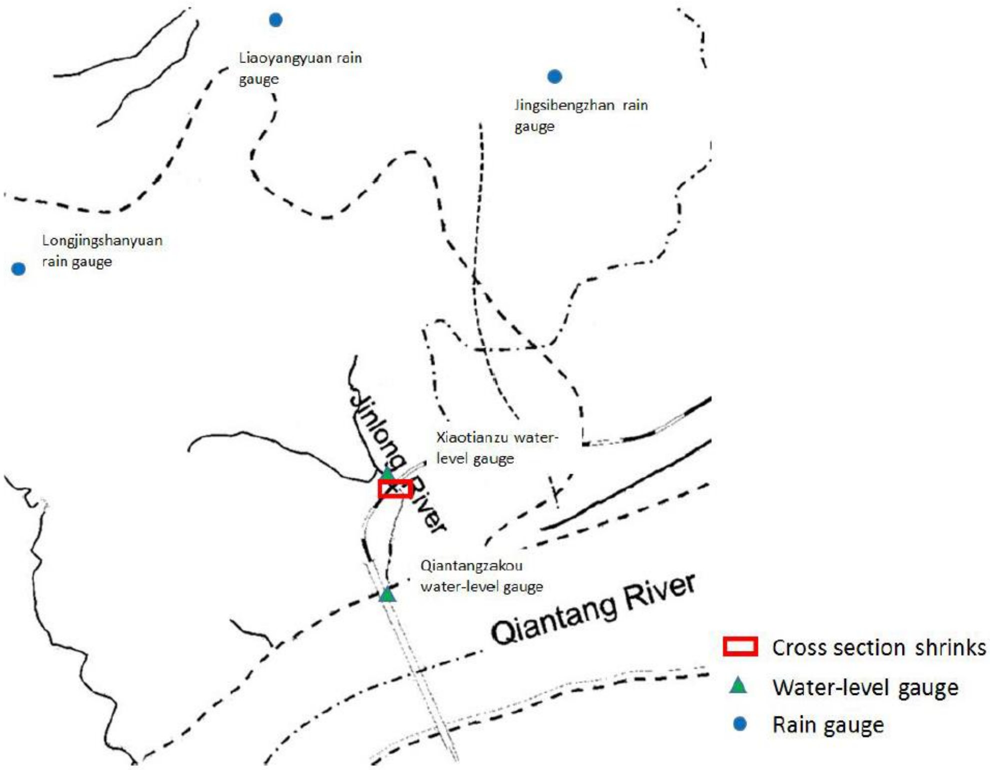

Fig. 1 Description of the study site

Framework of urban flood forecast

Framework of urban flood forecast in this study is shown in Fig. 2. It includes two components, i.e., MIKE FLOOD and SVM.

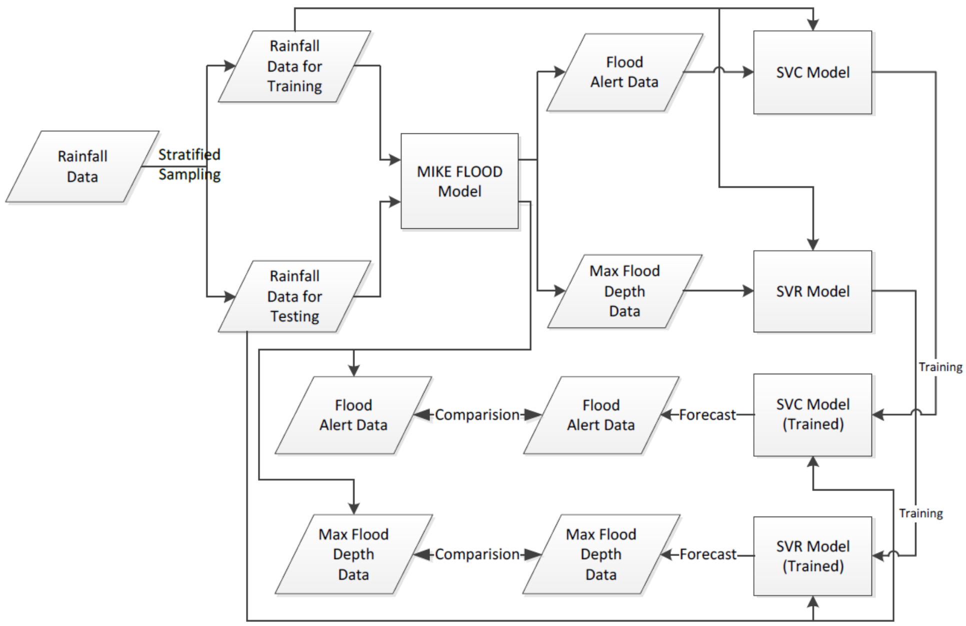

Fig. 2 Framework of urban flood forecast combining SVM and numerical model

MIKE FLOOD is a physical process based modeling suite, which has shown its great accuracy in simulating urban flood events (Hlodversdottir et al. 2015).SVM, a kind of machine learning algorithms, which is based on a structural risk minimization principle, has been verified to be a robust and efficient algorithm for equation fitting, data analysis and so on (Collobert and Bengio, 2001). Two SVM models were established in this paper: the maximum flood depth was forecasted by a support vector regress machine (SVR), and the urban flood alert was forecasted by a support vector classification machine (SVC). In this study, the urban flood alert was defined as flood depth above 0.15 meter lasts over 0.5 hour.

Both SVM models included training and testing steps. In the training step, the flood data extracted from MIKE FLOOD was used for training the SVM models. In the testing step, the MIKE FLOOD model results were used to assess the performance of the trained SVM models. Then the trained SVM models could be used in urban flood forecast.

Statistical Analysis of SVM models

Six performance measures comparing the flood results simulated by numerical model with the results simulated by SVM models were used to evaluate the performance of SVM models. Four parameters including root mean square error (RMSE), mean bias error (MBE), coefficient of efficiency (CE), and coefficient of correlation (CC), were used to evaluate the SVR model. Precise rate (PR) and true positive rate (TPR) were used to evaluate the flood alert forecasting accuracy forecasted by the SVC model.

Results and Discussion

MIKE FLOOD model calibration and verification

The water level data from Xiaotianzu gauging station of two storm events was used for model calibration (Oct. 6-8th, 2013) and verification (Jun. 26-29th, 2013). The model verification result (Fig. 3) showed a good similarity between the observed and MIKE FLOOD simulated result. At the rainfall peak, the absolute biases were below 0.05m.

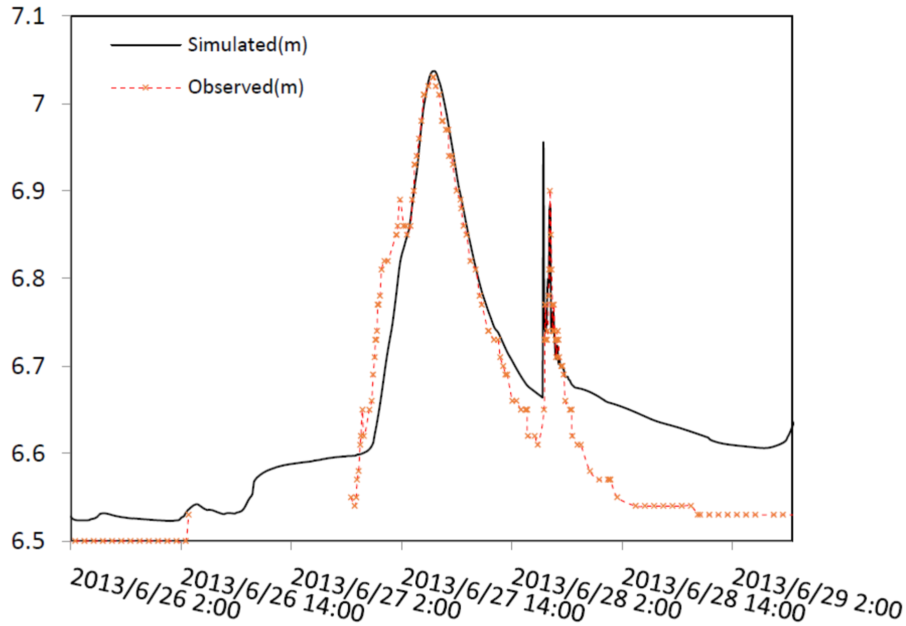

Fig. 3 MIKE FLOOD model verification (Jun. 26th-29th, 2013)

Performance of the SVM models

77 rainfall events were selected for training the SVM models and 34 rainfall events were selected for testing. The performance indicators are summarized in Tab. 1. Considering the testing events, PR is 96.9%, TPR is 92.9%, and the RMSE, MBE, CE and CC is 0.046, -0.011, 0.919 and 0.964 respectively, indicating the SVM models had similar forecast performance as the MIKE FLOOD model.

Tab. 1 Performance indicators of SVM models

	PR (%)	TPR (%)	RMSE	MBE	CE	CC
Training events	100	100	0.021	0.000	0.985	0.993
Testing events	96.9	92.9	0.038	-0.001	0.945	0.972

The MIKE FLOOD model and SVM models were run in the same desktop (Intel® Xeon® CPU E5-2687W v2, 32G RAM) in order to compare their computation time. In detail, for all the testing events, the total computation time based on MIKE FLOOD was 25 hours, while the SVM models only took 2.1 milliseconds.

Conclusions

To provide high effective real-time urban flood forecast, an urban flood forecast framework combined SVM model with a well calibrated numerical model based on MIKE FLOOD was developed.

An application in Jinglong River Basin, Hangzhou, China, demonstrated the superiority of the developed models. Among performance indicators used to check the performance of the SVM models, the PR and TPRH were all above 90%, and the CE and CC were all above 0.9. It showed the SVM models could provide accurate urban flood forecast, while offering 15 million times faster calculating speed compared with the numerical model.

In conclusion, combination of numerical model and SVM model will achieve high solution accuracy and save significantly computational time. The presented methodology demonstrates its potential as a valuable tool to provide high effective flash flood forecast and improve emergency responses to alleviate the loss of lives and property due to urban flooding.

References

Bach P. M., Rauch W., Mikkelsen, P. S., McCarthy, D. T., Deletic, A. 2014 A critical review of integrated urban water modeling-Urban drainage and beyond. Environmental Modelling and Software 54, 88-107.

Casulli, V. 2009 A high-resolution wetting and drying algorithm for free-surface hydrodynamics. International Journal for Numerical Methods in Fluids 60, 391–408.

DHI 2016 MIKE FLOOD 1D-2D modeling. User Manual. DHI Group, Horshølme, Denmark.

Collobert R., Bengio S. 2001 SVMTorch: Support vector machines for large-scale regression problems. The Journal of Machine Learning Research 1: 143-160.

Hlodversdottir A. O., Bjornsson B., Andradottir, H. O., Eliasson J., Crochet P. 2015 Assessment of flood hazard in a combined sewer system in Reykjavik city centre. Water Science and Technology 119, 1471-1477.