Model and Framework of Real-time Flood Process Detection under the Sensor Web Environment

Kridsakron Auynirundronkool
State Key Lab for Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, kron@gistda.or.th
*Wenying Du**
State Key Lab for Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, duwenying@whu.edu.cn
Nengcheng Chen
State Key Lab for Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, cnc@whu.edu.cn
Zeqiang Chen
State Key Lab for Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, zeqiangchen@whu.edu.cn

ABSTRACT

Flood occurred worldwide frequently, causing great causalities and economic losses. The occurrence and development of floods are process-based, while the state-of-art literature can only determine whether floods occur or not instead of judging flood phases, leading to coarse monitoring and delayed response. Therefore, it is imperative to construct the flood process detection (FPD) model and framework to implement the refined monitoring and provide decision information support for flood events. In this paper, floods were divided into four phases, including mitigation, preparedness, response, and recovery. Precipitation, water level, water flow and evaporation are variable indicators of floods, and their changes determine the flood occurrence or not in specific regions, thus the multi-dimensional vectors composed of precipitation, water level, water flow and evaporation are selected as flood indicators. The FPD model based on the support vector machine method was proposed in this paper to classify the flood indicators into four types, in correspondence with the four flood phases, respectively. The FPD framework conforming to the two standard information models, sensor model language, and observations & measurements, and the two service interfaces, sensor observation service, and sensor event service, of sensor web was proposed to achieve the real-time flood indicator access, filtering, and flood detection based on the aforementioned FPD model. The daily precipitation, water level, water flow, evaporation and the corresponding flood records data of Jiangxi, China from 1980 to 1989 were used as experimental data, firstly 80% of the data were used to construct the FPD model, and the left 20% were employed to verify the model precision, and secondly all the experimental observations were real-time accessed and

filtered to validate the feasibility of the framework. The high accuracy of the flood detection results demonstrated the feasibility of the FPD model and framework proposed in this paper.
Keywords- Flood Process Detection, Sensor Web, SVM, Real-time

1. INTRODUCTION

Flood occurred worldwide frequently, causing great causalities and economic losses. Thus, it is of great urgency to propose a flood detection method for early flood alert and disaster prevention [1]. There are two kinds of early flood discovery methods in the current literatures, including the long-term or medium-to-long term flood probability prediction (FPP) method [2], and the real-time or near real-time flood detection method (FDM) [3, 4]. The occurrences of historic flood events all conform to the rule of "flood 1, silence, flood 2, silence, flood 3, …", which is exactly the basis of FPP. FPP is to discover the variation law of observations during the process of flood occurrence and development by analyzing the long time-series flood records and their corresponding hydrologic and meteorological observations, and then make the probability of flood occurrence predictions based on the probability density distribution function.

Predictions of flood events were usually made based on the correlation analysis between flood record and its corresponding hydrological and meteorological observation. The variation law of observation data during floods were discovered in the correlation analysis, based on which floods in the future were predicted. The river stages, rainfall-runoff, flow and flash flood predictions based on neural networks [5-9] all belong to this kind of predictions. While these studies are all long-time predictions based on historic data and probability distribution rules, all with the characteristics of poor stability, large errors and not process-based.

Different from prediction, the FDM is to employ the critical conditions of flood occurrence or not to filter the real-time observations, and flood is thought of to happen when the conditions are met. Auynirundronkool et al. (2012) [10] tried to use the monitoring of the flood area changes to detect floods, and the moderate-resolution imaging spectroradiometer (MODIS) and Radarsat data were used in the flood detection and mapping during their experiments in central plains of Thailand. Martinis et al. (2014) [11] proposed a fully automatic flood area mapping service method, implementing the automatic process from satellite data download, preprocessing, to the flood area extraction. But all these methods are based on satellite images, while limited by spatial or temporal resolutions, it often occurs that there is no data available in the spatial and temporal range of floods. Besides, there are some flood detection system developed, including Jongman et al. (2015) [12], Khalaf et al. (2015) [13], Garcia et al. (2015) [14], Shi et al. (2015) [15], Lai et al. (2013) [16] but some of them are based on low-reliability or restricted data source, some of them using quite simple and idealized filter conditions, and others being not flexible and extensible.

In a word, the problems of flood detection are faced with can be summarized but not limited to the three points: (1) Lack of flood process detection; (2) Not real-time based; and (3) Inflexible and unextendible. Therefore, to resolve the three abovementioned problems, the model and framework for real-time flood process detection (FPD) were proposed in this paper. The FPD model was constructed based on the support vector machine (SVM) method [17], which is a classical and mature model for feature classification of satellite images, and here it was used in the classification of flood observation dataset. The FPD framework was built based on Strom, a streaming data processing framework, and the senor web information models as well as service interfaces, as the sensor web [18, 19] is recognized as the best candidate of implementing the unified while extensible real-time access, filter, and processing operations.

In the forthcoming sections, we illustrated the model and framework in Section II, performed the experiment and displayed the results in Section III. The discussion about the model and framework was provided in Section IV. Finally, Section V summarizes this work and describes future directions for this research.

2. METHODOLOGY

Real-time flood process detection under the sensor web environment was realized by the combination of the SVM-based FPD model and the real-time FPD framework. The former laid the model foundation for the determination of flood phases, and the latter provided a unified architecture for real-time data access, data filtering, and flood alert. The SVM-based FPD model was the core algorithm in the data filtering mechanism of the real-time FPD framework, and the real-time FPD framework provided a large-scale application scenario for the SVM-based FPD model.

2.1. SVM-based FPD model

The premise for the construction of SVM-based FPD model was that: (1) Floods were induced by the observable variables such as precipitation, and water level etc., for a specific region or watershed (with terrain, underlying surface, and drainage etc., remaining unchanged), and (2) Changes of the observable variables such as precipitation, and water level etc., during the occurrence and development of floods follow certain rules. The SVM-based FPD model was established based on the premise and the FLCNDEM proposed by Chen et al., 2015, in which floods are divided into four phases, including mitigation, preparedness, response, and recovery.

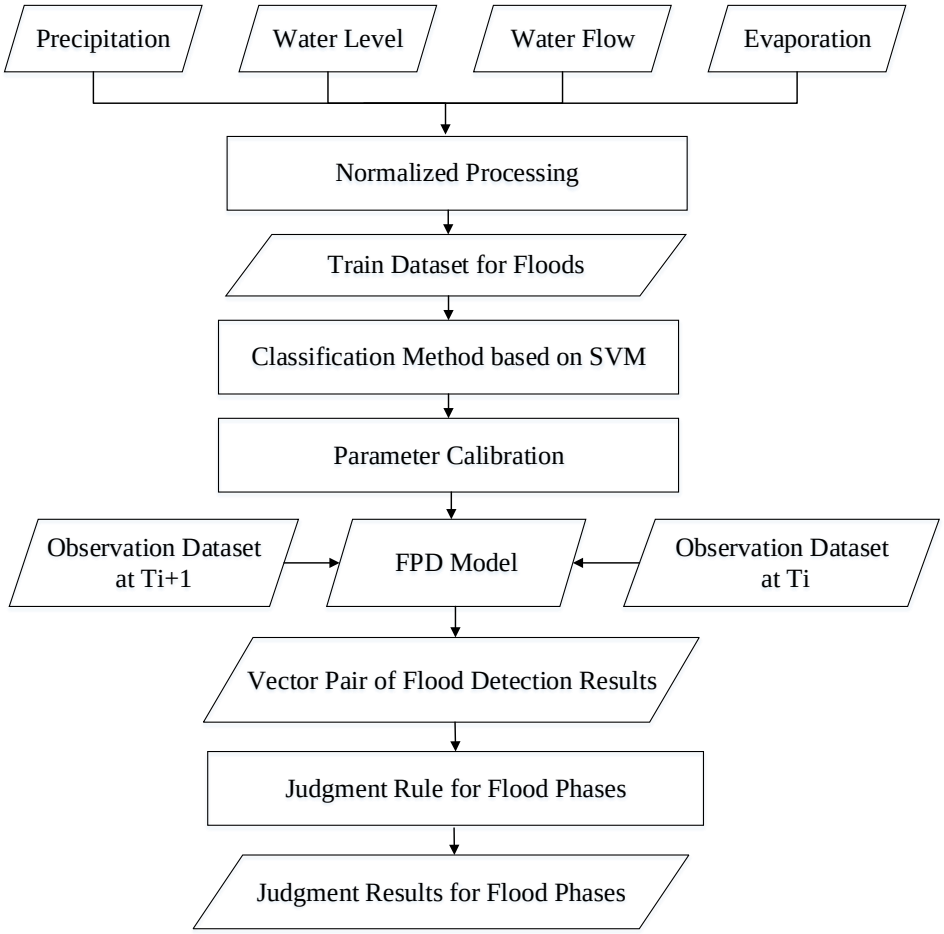

Fig. 1. Workflow of the SVM-based FPD model

In the SVM-based FPD model, as shown in Fig. 1, (1) the four observable variables of precipitation, water level, water flow, and evaporation were normalized according to the normalized observation calculation formula provided in Eq. (1); (2) each data record was composed of the normalized precipitation, water level, water flow, evaporation and their corresponding flood situation (FS) at the same day, FS = 1 represented there was flood while FS = 0 referred to there was no flood, and the FS values were determined by the authoritative flood records, all the data records consisted of the train dataset; (3) the FPD model was established based on the train dataset, the SVM model and the parameter calibration process; (4) Input the test observation dataset at the time T_i and T_{i+1}, and acquire the vector pair of flood detection results (FS_i, FS_{i+1}), where FS_i and FS_{i+1} refer to the FS at the time T_i and T_{i+1}, respectively; and (5) determine the flood phase according to the judgment rule demonstrated in Eq. (2).

$$\text{Normalized Observation} = \frac{\text{Observation - Average}}{\text{Max - Min}} \tag{1}$$

where observation refers to the original observation values of precipitation, water level, water flow, and evaporation, and average, max, and min refer to the average, max, and min value of them, respectively.

$$\text{Flood Phase} = \begin{cases} \text{Mitigation}, & \text{When}(FS_i, FS_{i+1}) = (0,0) \\ \text{Preparedness}, & \text{When}(FS_i, FS_{i+1}) = (0,1) \\ \text{Response}, & \text{When}(FS_i, FS_{i+1}) = (1,1) \\ \text{Recovery}, & \text{When}(FS_i, FS_{i+1}) = (1,0) \end{cases} \tag{2}$$

2.2. Real-time FPD Framework

The real-time FPD framework was able to implement the real-time observation data stream access, the FPD model based observation data stream filtering, the refined detection of flood phases, and E-mail alerts according to demands. It is composed of Storm, Sensor Observation Service (SOS), Sensor Event Service (SES), the connecting middleware of SOS and SES, SOS-SES-Feeder, and the SVM-based FPD model. Among these components, Storm and SOS takes charge of real-time observation data stream access, the SVM-based FPD model is used as the filtering algorithm of the observation data stream, SOS-SES-Feeder and SOS are to complete the data stream filtering and the flood phase determination.

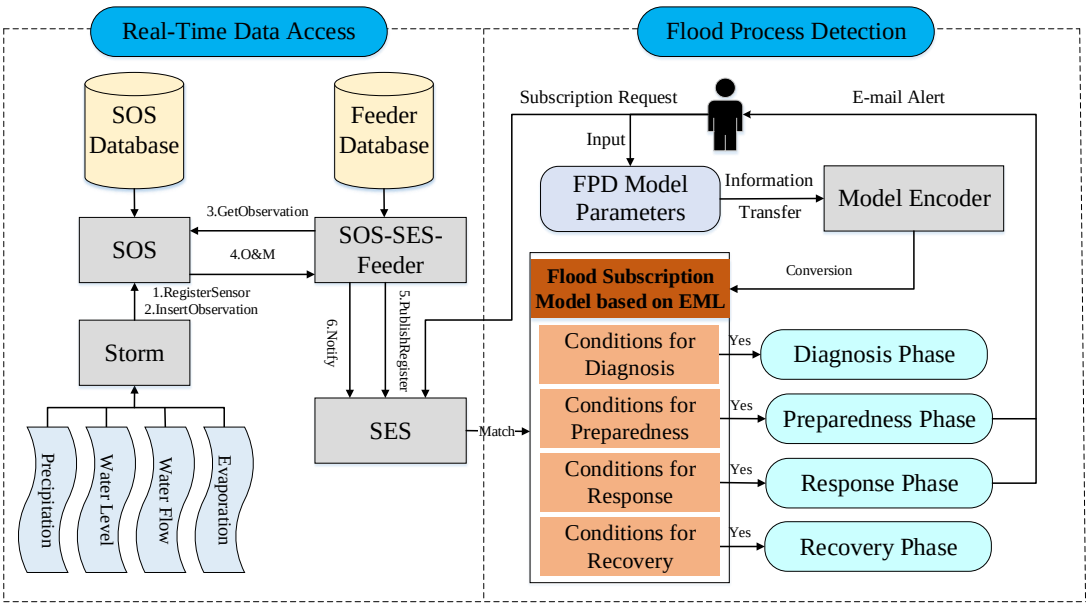

Fig. 2. Workflow of the Real-time FPD Framework

As demonstrated in Fig. 2, the workflow of the real-time FPD framework was: (1) Sensor Registration: the description information of the precipitation, water level, water flow, and evaporation sensors was encoded according to the Senor Model Language (SensorML), and was registered into SOS via the RegisterSensor operation; (2) Data Insertion: the observation data of the precipitation, water level, water flow, and evaporation sensors was encoded according to the Observations & Measurements (O&M), and was inserted into the SOS database through the InsertObservation operation; (3) Data Transfer: SOS-SES-Feeder sent the data request to SOS via the GetObservation operation, and SOS returns the data required to SOS-SES-Feeder in the format of O&M; (4) Data Registration in SES: SOS-SES-Feeder completed the PublisherRegistration operation and notify the registration results to SES; (5) User

Subscription: User submits the data subscription request and the FPD model parameters to SES, SES first received the parameters, encoded the model according to the EML format, formed the flood subscription model, then filtered the data stream according to the flood subscription model formed, and finally matched the data stream with the flood phase conditions, determined the flood phase and provided E-mail alert when necessary.

3. EXPERIMENTS AND RESULTS (14 Bold)

3.1. Experimental data

This paper used the precipitation, water level, water flow, and evaporation data in Jiangxi Province, China, from 1980 to 1989 (Download from the website of the National Earth System Science Data Sharing Infrastructure of China, URL: http://www.geodata.cn/data/), and the corresponding flood record data from the China meteorological disaster record collection (Jiangxi Volume) [20], and the flood records, precipitation, and water level observations of the Liangzi Lake flood occurring in Liangzi Lake, Wuhan, China from 1 July 2010 to 31 August 2010, as experimental data to test the validation and feasibility of the SVM-based FPD model and the real-time FPD framework, respectively. The reason why different dataset was used in the real-time FPD framework validation experiment was that the Liangzi Lake flood event lasted a longer time so that the results of different flood phases can be distinguished and displayed more clearly. The data frequency is one data record per day, and the flood records are specific to township. To be specific, there were 27 floods, totally lasting 68 days, covering 14 counties and 14 hydrological or meteorological stations, involved in our dataset. The space and time information of flood occurrence and their corresponding flood duration time and hydrological or meteorological station name information were displayed in Table 1. In contrast, non-flood days were also selected to consist of the dataset. There are in total 377 records in the dataset, 68 flood days (18%), 309 non-flood days (82%). Among the dataset, 299 records (79.3%) were used as the train dataset, and 78 records (20.7%) were used as the test dataset. For the train dataset, 54 records (18.1%) were flood days, while 245 records (81.9%) were non-flood days. And for the test dataset, 14 records (17.9%) were flood days, while 64 records (82.1%) were non-flood days. The summary of the record number and proportion for different types of dataset was shown in Table 2.

Table 1. Flood records and the corresponding stations

Location (County)	Time	Flood Duration (Days)	Station
	8 May 1983	1	
Shangyou	4-5 May 1984	2	Anhe
	2-3 July 1985	2	
Wanan	26 – 27 April 1980	2	Dongbei

	12 July 1980	1	
	14-15 May 1985	2	
	4 - 5 June 1985	2	
	26 July 1983	1	
Longnan	1 - 2 September 1984	2	Dutou
	3 July 1985	1	
Gaoan	14-18 June 1982	5	Jiacun
	7-9 July 1983	3	
Huichang	15-17 June 1983	3	Mazhou
Shanggao	7-9 July 1983	3	Niutoushan
	4 June 1985	1	
Lianhua	13-18 June 1982	6	Qianfang
	11-12 May 1983	2	
Wuning	27 June - 1 July 1981	5	Qingjiang
	30 May 1983	1	
Nanfeng	31 May - 1 June 1984	2	Shuangtian
Lean	13 - 18 June 1982	6	Wutou
	15 June 1983	1	
Xiushui	30 May 1983	1	Yangshuping
	5-10 July 1983	6	
Yifeng	6-9 July 1983	4	Yifeng
Yiyang	2 June 1983	1	Yiyang
Yongxiu	30-31 August 1980	2	Zhelin

Table 2. Summary of the record number and proportion for different types of dataset

	Total Records	**Flood Days**	**Flood Frequency**	**Non-Flood Days**	**Non-Flood Frequency**
Train Dataset	299	54	18.1%	245	81.9%
Test Dataset	78	14	17.9%	64	82.1%
Total Dataset	377	68	18.0%	309	82.0%

3.2. Detection results of the FPD Model

There are four parameters *s*, *t*, *c*, *g* involved in the FPD model, with *s* referring to the type of SVM, *t* representing the type of kernel function, *c* meaning loss function, and g referring to the gama function of the kernel function. In this paper, the parameter calibration process was performed by keeping one parameter changing while others remaining unchanged, and the parameter combination which could produce the highest overall accuracy of flood detection results were selected as the optimal parameters. The optimal parameters used in this paper were $s = 0$, $t = 2$, $c = 2$, and $g = 16$, respectively. Under the optimal parameter conditions, detection results of the FPD model were shown in Table 3, with 9 floods correctly detected while 5 floods not detected, and with 63 non-floods correctly detected while 1 non-floods mistakenly detected as flood.

Table 3. Detection results of the FPD Model

Class	Flood Record		Total
	Flood	Non-Flood	
Flood	9	5	14
Non-Flood	1	63	64

3.3 Real-time data access, filter, and alert

This section used the Liangzi Lake flood occurring from 1 July 2010 to 31 August 2010 as an example to test the feasibility of the real-time FPD framework. As the Liangzi Lake flood was a historic flood event, in order to test the real-time data access, filter, an alert, the data time was resampled to the system time of experiments, and the data frequency was resampled from one day to one minute to shorten the experimental time. A prototype was developed based on the real-time FPD model and framework proposed in this paper, and real-time data access interface was shown in Fig. 3. The real-time data access was realized by combing Storm and SOS, and what users need to do was to select the sensor you were interested in and clicked the finish sensor selection button, then the selected sensor would be registered into the system automatically, and its associated data could be accessed in real-time.

Fig. 3 Prototype interface for real-time data access

The result display interface for the Liangzi Lake flood was shown in Fig. 4, with different color representing different flood phases, cyan referring to the mitigation

phase, orange representing to the preparedness phase, pink meaning the response phase, and green referring to the recovery phase. The detection results conformed to the flood occurrence rule, that was, first mitigation phase, then preparedness and response phase, and finally the recovery phase.

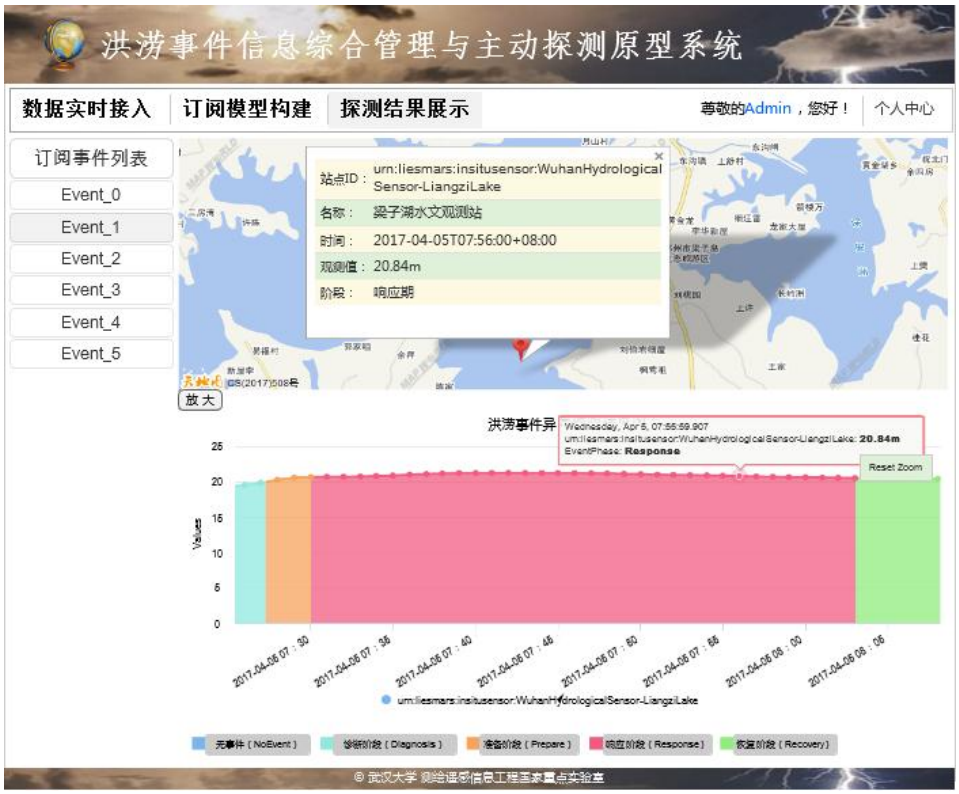

Fig. 4 Result display interface for the Liangzi Lake flood

Fig. 5 demonstrated the E-mail alert interface for the Liangzi Lake flood, (a) is the alert for the preparedness phase, while (b) is that for the response phase. Whether the E-mail alert was sent or not depended on the flood phase detection results. E-mail alerts would be sent to the subscriber only when the flood was detected to be in the preparedness or response phase.

(a) (b)

Fig. 5. E-mail alert interface for the Liangzi Lake flood

4. DISCUSSION

This section will be elaborated from the two perspectives of model accuracy evaluation as well as framework reusability and extensibility. The former performs quantitative evaluation for the SVM-based FPD model, and the latter discusses the reusability and extensibility of the real-time FPD framework.

4.1. Model Accuracy Evaluation

The accuracy of the FPD model determines the accuracy of flood detection. This section performs the quantitative accuracy evaluation for the SVM-based FPD models. The evaluation was performed based on the three quantitative indicators of overall accuracy (OA), Kappa Coefficient, Commission Error (CE), Omission Error (OE), Producer's Accuracy (PA), and User's Accuracy (UA) [21]. Based on the detection results provided in table 3, OA = 92.3%, CE = 7.1%, OE =35.7%, PA = 64.3%, and UA =92.9%, respectively.

4.2. Framework reusability and extensibility

The reusability and extensibility of the real-time FPD framework can be elaborated from the two perspectives of sensor providers and data consumers. Sensor providers publish their data through the framework, while data consumers use the data

published by sensor providers. Data exchange and sharing can be achieved through the real-time FPD framework proposed in this paper.

4.2.1. Service reusability and extensibility for sensor providers

The real-time FPD model and framework implemented in section 3 just provide an instance of the process-based flood detection and service. The framework can be reused and extended by sensor providers. For sensor providers, they can use the service to publish their sensor information and the sensor associated observations. If sensor providers have new sensors to be published, what they need do is just firstly encoding their sensor description information according to the SensorML template, and secondly formatting their observations in accordance with the O&M template.

4.2.2. Service reusability and extensibility for data consumers

For data consumers, they can reuse the real-time FPD model and framework to subscribe flood events occurring in different regions or different time ranges. For the subscription of flood events occurring in the same region but different time ranges, what the data consumers are required to do is just changing the time ranges of the data, and then submit their new subscriptions. But for the subscription of flood events occurring in the different regions and different time ranges, the data consumer need to change both the spatial and temporal ranges in their subscriptions. Besides, the SVM-based FPD detection model proposed in this paper can be extended to any model of flood detection, only if the detection model is encoded according to the specification of EML.

4.3. Advantages

The real-time FPD model and framework proposed in this paper provide a unified and extensible solution for process-based flood detection and alert, via which the flood events occurring in any region or any time all can be detected. New sensors, flood detection algorithms or flood services can be added in the real-time FPD model and framework according to the method provided in section 4.2.1 and 4.2.2. The real-time FPD model and framework makes a flood detection model and system universally applied in multiple regions and organizations possible, and it can increase the reuse rate of the code, meanwhile save plenty of time and resources possibly wasted by repetitive coding.

4.4. Limitations

Parameters involved in the SVM-based FPD model was calculated based on the historic flood associated observations, and inappropriate parameters might lead to inaccurate flood detection results. In addition, the real-time FPD framework proposed in this paper could only provide alert service, and more service types need to be added in the framework to satisfy the needs from different phases of floods.

5. CONCLUSION AND OUTLOOK

In this paper, firstly, in combination with flood observation data and flood records, the SVM-based FPD model was proposed in this paper; secondly, the real-time FPD framework was put forward based on Storm, SOS, SES, and the SVM-based FPD model; and thirdly, the flood event process detection and alert prototype is designed and implemented based on the SVM-based FPD model and the real-time FPD framework. The water level, precipitation, water flow, and evaporation data of ** hydrological stations in Jiangxi, China from year 1980 to 1989 and their corresponding flood record data were used as experimental data to test the feasibility and effectiveness of the real-time FPD model and framework proposed in this paper, and the detection results demonstrated that the model and framework proposed in this paper could be used in the detection of flood process, proving that the feasibility and effectiveness of the SVM-based model and the real-time FPD framework.

ACKNOWLEDGEMENTS

This work was supported by grants from the National Nature Science Foundation of China program (no. 41771422), the Nature Science Foundation of Hubei Province (no. 2017CFB616), the Union Foundation of Ministry of Education of the People's Republic of China (no. 6141A02022318), the Creative Research Groups of Natural Science Foundation of Hubei Province of China (no. 2016CFA003) and the Fundamental Research Funds for the Central Universities (no. 2042017GF0057).

REFERENCES

[1] GuhaSapir, et al., "Annual disaster statistical review 2014: the numbers and trends," Region II - Asia, 2015.

[2] Saini, S.S., S.P. Kaushik, and R. Jangra, "Flood-risk assessment in urban environment by geospatial approach: a case study of Ambala City, India," Applied Geomatics, 8(3): pp. 1-28, 2016.

[3] Cornelia, G. and R. Peter, "Multitemporal and Multispectral Remote Sensing Approach for Flood Detection in the Elbe-Mulde Region 2002," CLEAN - Soil, Air, Water, 33(5): pp. 395-403, 2010.

[4] Rudolfbrãzdil, Z. Kundzewicz, and Gerardobenito, "Historical hydrology for studying flood risk in Europe," International Association of Scientific Hydrology Bulletin, 51(5): pp. 739-764, 2006.

[5] Alvisi, S. and M. Franchini, "Grey neural networks for river stage forecasting with uncertainty," Physics & Chemistry of the Earth, 42-44(208): pp. 108-118, 2012.

[6] D. Herr, H. and R. Krzysztofowicz, "Addendum to Bayesian ensemble forecast of river stages and ensemble size requirements,". 387: pp. 151-164, 2010.

[7] Khan, M.S. and P. Coulibaly, "Bayesian neural network for rainfall-runoff modeling," Water Resources Research, 420(7): pp. 379-393, 2006.

[8] Piotrowski, A., J.J. Napiórkowski, and P.M. Rowiński, “Flash-flood forecasting by means of neural networks and nearest neighbour approach – a comparative study,” Nonlinear Processes in Geophysics, 13(4): pp. 443-448, 2006.
[9] Sahoo, G.B. and C. Ray, “Flow forecasting for a Hawaii stream using rating curves and neural networks,” Journal of Hydrology, 317(1–2): pp. 63-80, 2006.
[10] Auynirundronkool, K., et al., “Flood detection and mapping of the Thailand Central plain using RADARSAT and MODIS under a sensor web environment,” International Journal of Applied Earth Observations & Geoinformation, 14(1): pp. 245-255, 2012.
[11] Martinis, S., J. Kersten, and A. Twele, “A fully automated TerraSAR-X based flood service,”. ISPRS Journal of Photogrammetry & Remote Sensing, 104(2013): pp. 203-212, 2015.
[12] Jongman, B., et al., “Early Flood Detection for Rapid Humanitarian Response: Harnessing Near Real-Time Satellite and Twitter Signals,”. ISPRS International Journal of Geo-Information, 4(4): pp. 2246-2266, 2015.
[13] Khalaf, M., et al. “Advance flood detection and notification system based on sensor technology and machine learning algorithm,” International Conference on Systems, Signals and Image Processing. 2015.
[14] Garcia, F.C.C., A.E. Retamar, and J.C. Javier. “A real time urban flood monitoring system for metro Manila,” in TENCON 2015 - 2015 IEEE Region 10 Conference. 2016.
[15] Shi, H., et al., “A service-oriented architecture for ensemble flood forecast from numerical weather prediction,” Journal of Hydrology, 527: pp. 933-942, 2015.
[16] Lai, H.C., et al., “A Developing Flood Automation System,” Applied Mechanics & Materials, 457-458: pp. 1220-1223, 2013.
[17] Cortes, C. and V. Vapnik, “Support Vector Network,” 20(3): pp. 273-297, 1995.
[18] Botts, M., et al. “OGC® Sensor Web Enablement: Overview and High Level Architecture,” in Autotestcon. 2008.
[19] Arne, B., et al., “New Generation Sensor Web Enablement,” Sensors, 11(3): pp. 2652-99, 2011.
[20] Wen, K., China Meteorological Disaster Records (Jiangxi Volume) (In Chinese). Beijing, China: Meteorological Press, 2006.
[21] Fitzgerald, R.W. and B.G. Lee, “Assessing the classification accuracy of multisource remote sensing data,” Remote Sensing of Environment, 47(3): pp. 362-368, 1994.