Quantifying the contribution of flood intensity indicators with the projection pursuit model

Lina Wang, Yanqing Lian and Xiaohong Chen

ABSTRACT

Identifying the various factors that affect the intensity of a flood event, such as its duration and volume, is essential for strategic planning and flood management. Further, quantifying the impacts of these major factors on flood intensity using the contribution rate is essential, but technically challenging. In this study, the authors have adopted the projection pursuit model to quantify the contribution rates of peak flood stage and peak flood discharge, flood duration, and total flood volume (the maximum 12-, 24-, and 72-hour flood volumes) in the Wujiang River in Southern China. This study showed that peak flood discharge and total flood volume were the two dominant factors impacting flood intensity. Although flood duration can be a major factor for some flood events, it contributed the least to flood intensity for most of the historic flood events studied. Likewise, the maximum 24-hour and 72-hour flood volumes contributed little to flood intensity. Findings from this study not only demonstrated the successful adoption of the projection pursuit model for contribution rates, but also provided critical information for planning and managing the regional hydraulic resources in the Wujiang River.

Key words | contribution rate, flood intensity, key driving factors

Lina Wang
School of Geography,
South China Normal University,
Guangzhou 510631,
China

Yanqing Lian (corresponding author)
Prairie Research Institute,
University of Illinois at Urbana-Champaign,
2204 Griffith Drive, Champaign, IL 61820,
USA
E-mail: *ylian@illinois.edu*

Xiaohong Chen
Center for Water Resources and Environment,
Sun Yat-sen University,
Guangzhou 510275,
China
and
Key Laboratory of Water Cycle and Water Security in Southern China of Guangdong Higher Education Institutes,
Sun Yat-sen University,
Guangzhou 510275,
China

INTRODUCTION

Floods can be considered catastrophes, causing lost lives and properties, and natural environments that are damaged beyond repair (Kisss *et al.* 2014). The severity of damage from a flood event can be assessed by determining the flood intensity. Factors that contribute to the intensity include the peak flood discharge, volume and duration, and flow velocity. Ahmed & Mirza (2000) suggested using the flood intensity index to account for multiple factors such as the flood duration, destructive capacity, and the area affected by floods. Many studies have examined flood intensity from different perspectives. For example, Adamowski (2000) pointed out that the peak flood discharge is the most important feature of a flood event because the maximum discharge rather than the average discharge is the main concern. French & Miller (2011) stated that the flood intensity is determined by flow depth and velocity. Swades & Surajit (2012) used three variables, including flood stage, flood frequency, and the flood stagnation period, to evaluate the flood intensity. Wang *et al.* (2015) proposed using different indicators to rank flood intensity. Javelle *et al.* (2003) pointed out that flood severity is not only defined by its peak value, but also by its volume and duration. An extensive amount of research has examined the impact of climate factors on flood intensity (Droegemeier *et al.* 2000; Changnon *et al.* 2001; O'Connor & Costa 2003, 2004). Flood volume may have the largest impact on flood intensity for large river basins, although the impacts of multiple factors on flood intensity would vary with respect to hydrologic, hydraulic, and climate conditions.

doi: 10.2166/nh.2017.038

While peak flood discharge, peak stage, and flood volume are regarded as the three major contributors to flood intensity, for practical purposes the peak flood discharge and its corresponding stage are often used to quantify the flood intensity (Jumadar *et al.* 2008; Singh & Singh 2011; Swades & Surajit 2012; Ward & Paulus 2013). The impacts of peak flood discharge and discharge on a flood event can be quantified by using contribution rates, or the weight of the variables that contributed to flood intensity. The contribution rate is an important, comprehensive value used for flood control. Quantifying contribution rates of factors that could significantly impact flood intensity not only provides information to identify the dominating factors for flood events, but also is crucial for flood risk assessment and water resource planning (Bai *et al.* 2015). The evaluation of a contribution rate should be as accurate as possible given the implications for designing hydraulic structures, water management, and estimation of scour at a hydraulic structure (Browne & Hoyt 2000; Carolan 2007; Shabri *et al.* 2014).

Although the principal component analysis (PCA) model has been used to calculate the contribution rate (Grimalt *et al.* 1990; Domingo *et al.* 2000; Kharmouz *et al.* 2013; Tomašić *et al.* 2014; Miloudi *et al.* 2017), the PCA method itself can only quantify the contribution rate of principal components. The principal components often consist of several variables, thus it is difficult to determine the contribution rate of a single or a specific variable (Hu *et al.* 2007; Noori *et al.* 2010, 2011). The projection pursuit algorithm is a statistical technique for finding the most 'interesting' possible projections in multidimensional data. The idea of projection pursuit is to locate the projection or projections from high-dimensional space to low-dimensional space that reveal the most details about the structure of the data set. It has been widely used in many fields such as multivariable prediction, cluster, water resources assessment, environmental protection, etc. (Kennedy & Basu 2000; Wang *et al.* 2004; Rajeevan *et al.* 2007; Chi & Dong 2009; Sun *et al.* 2012; Balke *et al.* 2013; Zhao *et al.* 2014; Jie *et al.* 2015; Durocher *et al.* 2016). Wang & Zhang (2010) developed the projection pursuit dynamic cluster model, which is based on the projection pursuit principle, to analyze the carrying capacity of water resources in China. Huang & Zhang (2011) proposed a flood disaster classification assessment method, which adopted the multi-swarm system particle swarm optimization method to optimize the parameters of the projection index functions. Liu *et al.* (2016) evaluated water resource resilience based on a projection pursuit classification model, with the help of artificial fish-swarm algorithm (AFS), which optimizes the projection index function.

The main objective of this study was to adopt the projection pursuit model to quantify contributions from multiple indicators, such as peak level, peak flood discharge, flood volume, flood duration, and the maximum t-hour volume. This study used historic flood event data to demonstrate the application of this method in the Wujiang River.

STUDY AREA AND DATA

The Wujiang River is one of the largest tributaries of the Beijiang River in the Pearl River basin in southern China. The Wujiang River basin is located south of the Wuling mountains between the latitudes of 24°46′ to 25°41′ N and longitudes of 112°23′ to 113°36′ E (Figure 1). Its total drainage area is about 7,097 km^2 with a mainstream spanning 260 km. The climate of the Wujiang River is dominated by the southwest and southeast Asian monsoons, which results in high humidity in the summer and an uneven distribution of precipitation throughout the year. The mean annual rainfall within the basin is approximately 1,450 mm. The Lishi stream-gaging station is located near the mouth of the river, monitoring conditions for 98.2% of the drainage area, or 6,976 km^2 of the Wujiang River basin (Figure 1). The stage and flow data from 1955 to 2007 at the Lishi station were provided by the Shaoguan Branch of the Guangdong Provincial Bureau of Hydrology. According to this agency, the discharge data were computed from the direct flow velocity measurements at the gaging station and the water level was measured by using a stage gage at the same location. The flood event with the largest peak flood discharge in each calendar year was selected for the study analysis. The method for extracting flood events was introduced in detail by Wang *et al.* (2015).

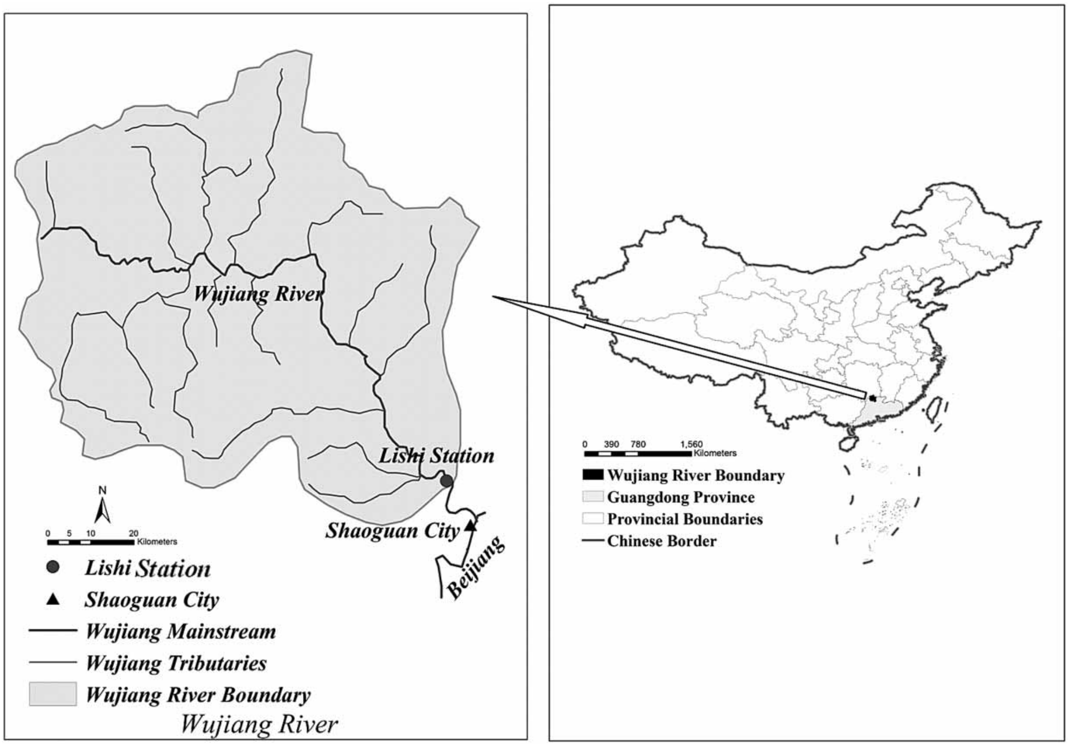

Figure 1 | Location of the study region and hydrological stations.

METHODOLOGIES

Indicators of flood events

The peak flood discharge, peak flood stage, total flood volume, and flood duration are often used as indicators for flood intensity. In addition, the maximum 12-h, 24-h, and 72-h flood volumes were also used as major indicators in this study. Wang *et al.* (2015) provided calculation details for the flood duration, the maximum 24-h and 72-h flood volume, the total flood volume, and other factors. Wang *et al.* (2015) showed that the average flood duration in the Wujiang River basin was 3 days.

Calculation of contributing rate

The projection pursuit model is designed to determine the projection from high-dimensional spaces to low-dimensional spaces that reveal the most details about the structure of the data set. Details of the projection pursuit model for the contribution rate of indicators can be found in Friedman & Tukey (1974). The steps of the projection pursuit model for contribution rate of indicators are as follows (Friedman & Tukey 1974; Jin *et al.* 2005; Huang & Zhang 2011; Zheng & Lin 2012).

The projection pursuit model is made to convert the p-dimensional data $\{x_{ij}|j = 1, 2, \cdots, p\}$ for $i = 1, 2, \cdots, m$ into one-dimensional data (Z_i), referred to as the projection value with the projection direction $a' = (a_1, a_2, \cdots, a_p)$ (Qin & Lin 2012). According to the projection component value (Z_{ij}) of the indicator j-th of i-th flood process, the indicator contribution rate can be expressed by Z_{ij}/Z_i, where Z_i synthesizes the projection value, and Z_{ij} is the projection component value receptivity (Wang *et al.* 2004).

Step 1. Establishment of indicator

Given the matrix $X^0_{mp} = \begin{bmatrix} x^0_{11} & x^0_{12} & \cdots & x^0_{1p} \\ x^0_{21} & x^0_{22} & \cdots & x^0_{21} \\ \cdots & \cdots & \cdots x^0_{ij} \cdot & \cdots \\ x^0_{m1} & x^0_{m2} & \cdots & x^0_{mp} \end{bmatrix}$ present the indicator values of flood process samples, where p is the number of flood process indicator variables, m is the number of flood events analyzed, and x_{ij}^0 denotes the value of the indicator j-th of i-th flood process. In this study, $m = 53$, $p = 7$ and

$$x_{i1}^0 = \{P_{11}, P_{21}, \cdots, P_{i1}, \cdots, P_{m1}\},$$

$$x_{i2}^0 = \{H_{12}, H_{22}, \cdots, H_{i1}, \cdots, H_{m2}\},$$

$$x_{i3}^0 = \{(V_{12})_{13}, (V_{12})_{23}, \cdots, (V_{12})_{i3}, \cdots, (V_{12})_{m3}\},$$

$$x_{i4}^0 = \{(V_{24})_{14}, (V_{24})_{24}, \cdots, (V_{24})_{i4}, \cdots, (V_{24})_{m4}\},$$

$$x_{i5}^0 = \{(V_{72})_{15}, (V_{72})_{25}, \cdots, (V_{72})_{i5}, \cdots, (V_{72})_{m5}\},$$

$$x_{i6}^0 = \{V_{16}, V_{26}, \cdots, V_{i6}, \cdots, V_{m6}\},$$

$$x_{i7}^0 = \{T_{17}, T_{27}, \cdots, T_{i7}, \cdots, T_{m7}\}.$$

Step 2. Normalization of data

Units for the seven variables selected in this study are different. Four types of data pretreatment methods including the mean centering, the differentiation, normalization, and auto-scaling can be used to eliminate dimensions of different process variables (Amrhein *et al.* 1996). The normalization method was adopted in this study and is given below:

$$x_{ij} = \frac{x_{ij}^0 - x_{j\min}^0}{x_{j\max}^0 - x_{j\min}^0} \tag{1}$$

where $X_{j\min}$ and $X_{j\max}$ are the initial minimum and maximum values of the j-th indicator, respectively.

After normalization, the matrix $X_{mp}^0 = \begin{bmatrix} x_{11}^0 & x_{12}^0 & \cdots & x_{1p}^0 \\ x_{21}^0 & x_{22}^0 & \cdots & x_{21}^0 \\ \cdots & \cdots & \cdot x_{ij}^0 \cdot & \cdots \\ x_{m1}^0 & x_{m2}^0 & \cdots & x_{mp}^0 \end{bmatrix}$ could be replaced by matrix

$$X_{mp} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \cdots & \cdots & \cdot x_{ij} \cdot & \cdots \\ x_{m1} & x_{m2} & \cdots & x_{mp} \end{bmatrix}.$$

Step 3. Analysis of linear projection

The p-dimensional data (X_{mp}) can be converted into one-dimensional data (Z_i) of the projection direction by the following formula:

$$Z_i = \sum_{j=1}^{p} a_j x_{ij} \quad (i = 1, 2, \cdots, m, j = 1, 2, \cdots, p) \tag{2}$$

where Z_i is the synthesizing projection value.

Step 4. Establishment of objective function

The projection index function can be constructed as follows:

$$Q_a = S_a \cdot d_a \tag{3}$$

where S_a, which equals the standard variance of the projection value Z_i, is the distance function between clusters. d_a is the density function in clusters, which means that the local density of the projection value is Z_i.

$$\begin{aligned} S_a &= \sqrt{\frac{1}{m-1}\sum_{i=1}^{m}(Z_i - \bar{Z})^2} \\ d_a &= \sum_{i=1}^{m}\sum_{k=1}^{m}(R - r_{ik}) f(R - r_{ik}), \ (i, k = 1, 2, \cdots, m) \end{aligned} \tag{4}$$

where $\bar{Z} = 1/m \sum_{i=1}^{m} Z_i$, $R = 0.1S_a$. R is the window radius of density function in clusters, which is related to data characteristics, while the selection of an average number of window projection points should not be too small, and the average deviation of glide should not be too large. *rik* is the absolute distance between the random projection eigenvalues, and $r_{ik} = |Z_i - Z_k|$. $f(R - r_{ik})$ is the unit step function, when $R > r_{ik}$, $f(R - r_{ik}) = 1$, or vice versa, $f(R - r_{ik}) = 0$.

Step 5. Optimization of objective function

The projection index function proposed by Friedman & Tukey (1974) can be written as:

$$\begin{cases} \max(Q_a) = S_a \cdot d_a \\ \sum_{j=1}^{p} a_j^2 = 1 \end{cases} \tag{5}$$

where a is the projection axis vector, also called the projection vector.

In this study the artificial fish-swarm algorithm (AFS) was adopted to find the optimizing value. AFS was used to solve the non-linear optimization problem of the projection pursuit model. An important difference between AFS and other swarm intelligence algorithms is that the AFS can search for the global optimum effectively and has an adaptive ability for the search space. According to Yazdani *et al.* (2010) and Mehdi *et al.* (2012), the AFS algorithm is one of the best optimization methods among the swarm intelligence algorithms. The AFS is a simulation behavior and population-based optimization method, which was initially developed by Tu (1999) and has recently been widely used in practice (Yazdani *et al.* 2012; Reza & Bojnourd 2014).

Step 6. Calculation of contribution rate

The contribution rate can then be computed by using the following expression:

$$C_{ij} = \frac{Z_{ij}}{Z_i} = \frac{a_j x_{ij}}{Z_i}, \; (i = 1, 2, \cdots, m; j = 1, 2, \cdots, p) \tag{6}$$

in which, $a_j > 0$, $(j = 1, 2, \cdots, p)$ to make sure that all of the optimization projection directions are positive.

RESULTS AND DISCUSSION

Flood intensity

Shown in Figure 2 are the relationships among the seven variables or indicators selected in this study. It can be seen, in general, that the flood peak discharges correlate well with flood peak stage, the 12-hour and the 36-hour flood volumes but not very well with the 72-hour flood volume and not well with the total flood volume and flood duration. Peak flood stages also correlate well with the 12-hour and 36-hour flood volumes but not well with the 72-hour and total flood volumes and not with the flood duration either. Examination of data show that 1975 was the seventh largest flood event according to the peak flood discharge but was the tenth largest flood in terms of the flood peak stage and 1993 flood was the eighth based on the peak flood discharge but was the seventh in terms of the flood peak stage. Figure 2 and data examination show that flood intensity is characterized by multiple facts and further justifies the use of several variables in this analysis.

In order to facilitate comparison and understand the difference between different flood events, cluster analysis has been used to compare the intensity of various flood events (Bhaskar & O'Connor 1989; Dong *et al.* 2007; Cheng *et al.* 2009; Latt *et al.* 2015; Wang *et al.* 2015). Clustering is a method for dividing scattered groups of data into several groups. Thus, the object of flood clustering is to sort flood into groups, so that the flood characteristic is similar between members of the same cluster and dissimilar between members of different clusters. Clustering analysis of flood intensity aims at the identification of groups of floods with common characteristics. Factors that affect flood intensity contain different units. Therefore, it is impossible to compare them by using their absolute values. Instead, dimensionless values were processed using Equation (1), introduced earlier for the cluster analysis.

Initial values need to be set in order to compute the projection value of each flood event. In this study, the initial population size was set at 30, the artificial fish perception scope was visible as 0.3, the largest number of tests of each move try number was 20, the crowd fact was 0.3, the maximum number of iterations was 60, and the moving step length was 0.1. The optimization projection direction a, obtained from the AFS method, would be:

$$a = (0.369 \; 0.444 \; 0.398 \; 0.2830 \; 0.301 \; 0.565 \; 0.122).$$

From the optimization results, the projected characteristic value (Z_i) and the cluster results can be calculated. Larger projection values indicate higher flood intensities.

Figure 3 shows that the projection results matched very well with actual floods in the Wujiang River. Analyses from the projection pursuit model also indicated that the 2006 flood was the largest flood ever recorded in the Wujiang River, which was consistent with the previous analysis by Wang *et al.* (2015).

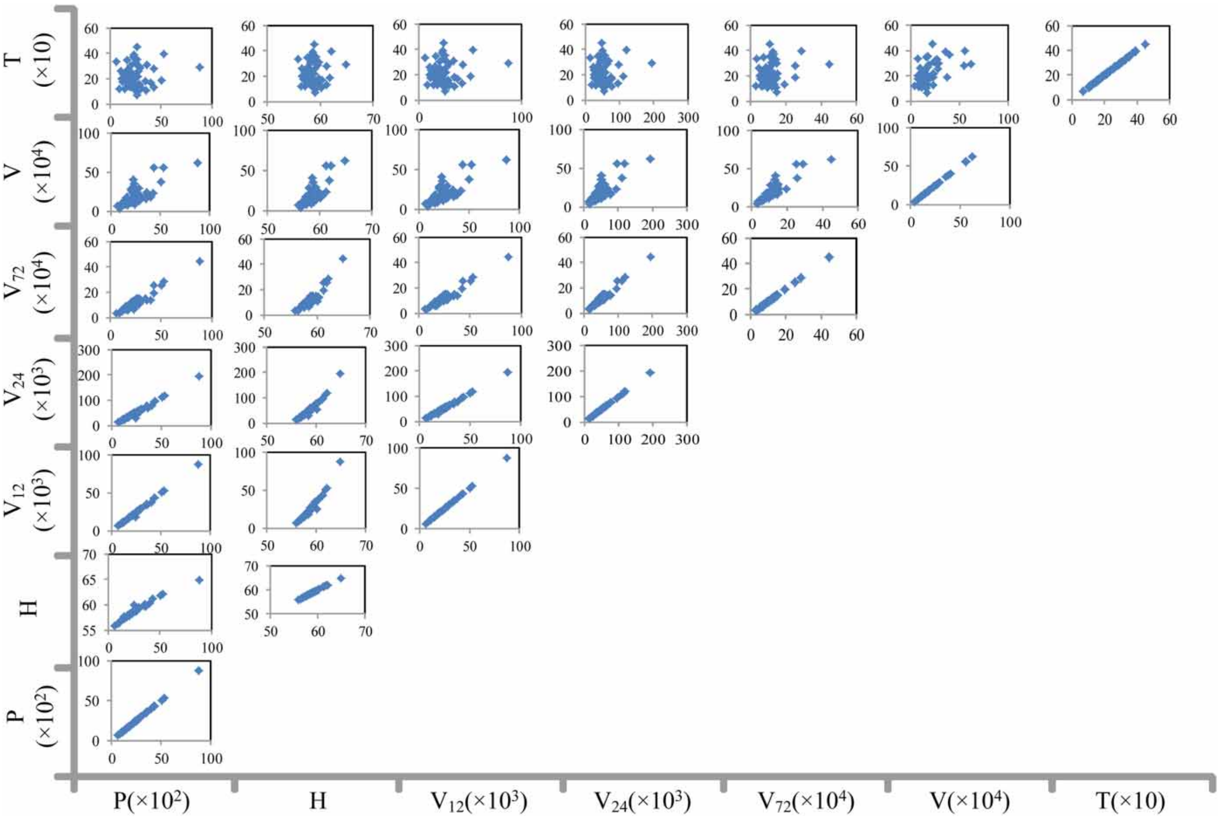

Figure 2 | Scatter plot of the variable matrix between different flood indicators. P: peak discharge (m^3/s), H: peak flood stage (m), V_{12}: maximum 12-h flood volume (m^3), V_{24}: maximum 24-h flood volume (m^3), V_{72}: maximum 72-h flood volume (m^3), V: total flood volume (m^3), T: flood duration (h).

Contribution rate

The calculation of contribution rate proposed in this study was the ratio of projection value for indicator to the projection value of all indicators. The project value for a single indicator is the product of projection direction and its normalized value of that indicator. The optimized values for the projection direction showed that the peak flow stage and the total flood volume had relative larger values which ultimately resulted in higher contribution rates for both among the seven indicators used in the analysis. Figure 4 shows the contribution rates of these seven impact factors in the 53-year period. It is apparent that the contribution rates for the same indicator were different among different flood events. The annual contribution rate for the peak flood discharge, peak stage, maximum 12-h, 24-h, and 72-h volume showed a slight increasing trend, whereas the annual contribution rate for the total flood volume and flood duration presented a slight decreasing trend, as shown in Figure 4(f) and 4(g), respectively.

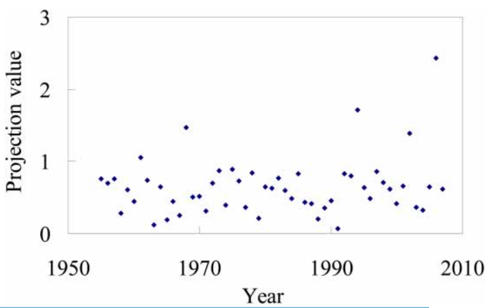

Figure 3 | The projection results of seven indicators from 1955 to 2007.

Figure 5 shows that the peak flood discharge stage was consistently one of the two major factors for all years except for 1963. The total volume of floods was the other major contributor for all years other than 1963, 1979, and

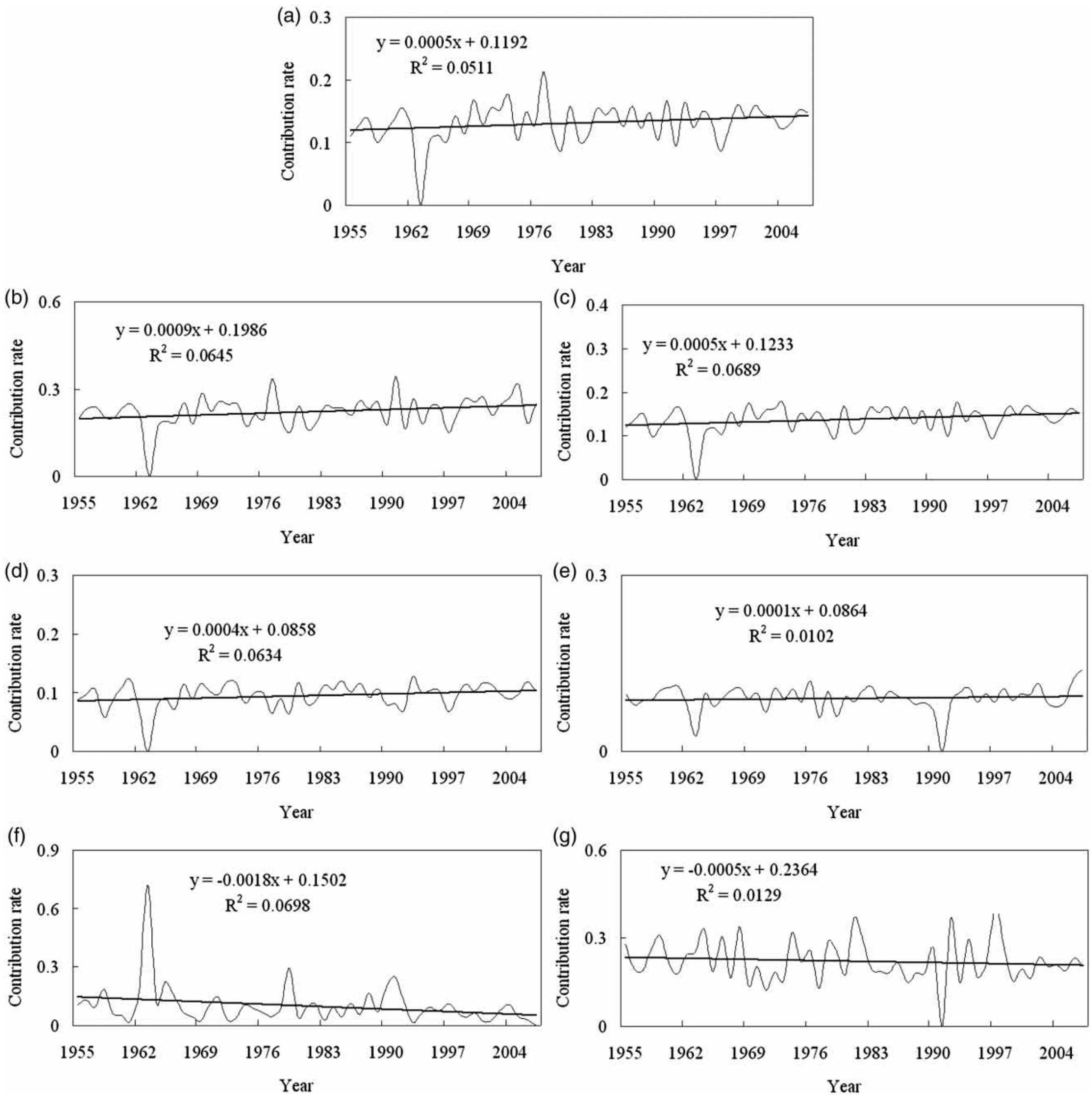

Figure 4 | Time series of indicator contribution rate at Lishi Station of the Wujiang River. Solid curve: the contribution rate of seven indicators from 1955–2007; solid line: the linear trends. The graph includes the thick straight-line pattern equation and the value of the determination index R^2. (a) Peak flood discharge; (b) peak level; (c) the maximum 12-h volume; (d) the maximum 24-h volume; (e) the maximum 72-h volume; (f) the total flood volume; (g) flood duration.

1991, during which the duration of flood events weighed more than the total flood volume. The extended length of the flood event was the major cause of flooding in 1963. The 1963 flood occurred in late March and early April, with a lowest peak flood discharge of 630 m^3/s. While this study did not show that the peak flood discharge was the major contributing factor or equally as important as the peak stage, the total discharge of a flood event was consistently the major contributor.

Characteristics of contribution rates

Figure 6 shows the indicators and their maximum contribution rates in each calendar year. Peak flood stage had the largest contribution rate for 29 of the 53 flood events, or 55% of the flood events. Total flood volume had the largest contribution rates for 21 of 53, or 40% of the flood events. In about 6% of the years, the flood duration had the largest contribution rates. Peak flood discharge, the

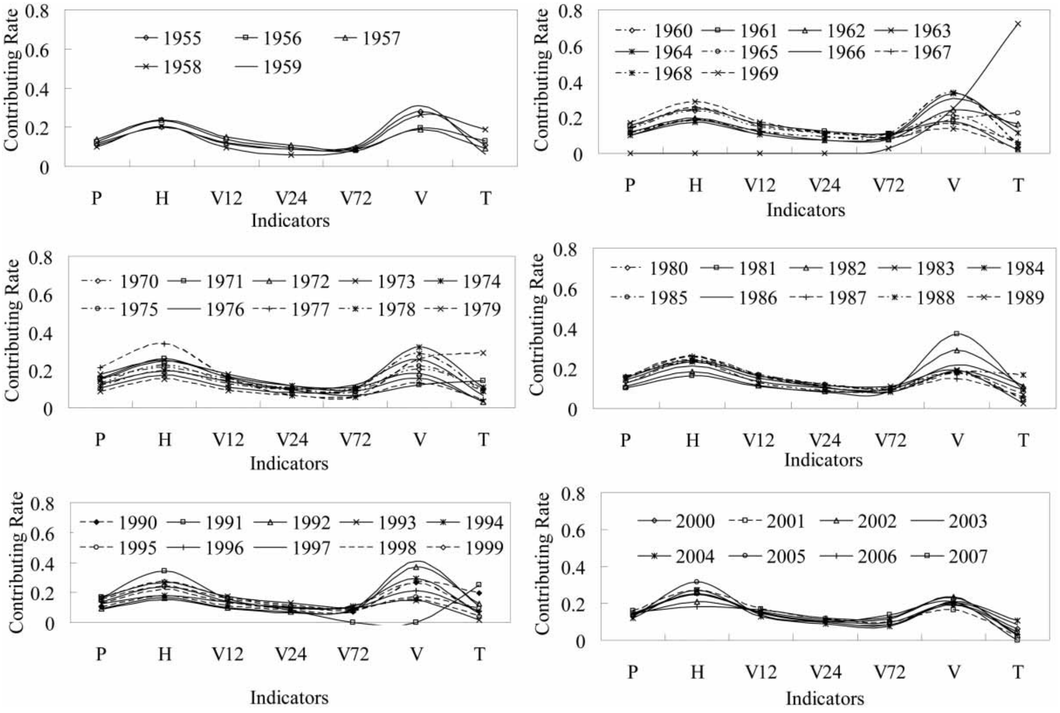

Figure 5 | The characteristics of the contribution rate. P: peak flood discharge; H: peak level; V12: the maximum 12-h volume; V24: the maximum 24-h volume; V72: the maximum 72-h volume; V: the total flood volume; T: flood duration.

maximum 12-h, 24-h, and 72-h volume never had the largest contribution rate in all 53 years.

Although interest and concern are often focused on factors that contribute the most to flood events, it is still useful to determine the factors that have had the least impact on flood events. Figure 7 shows that the flood duration had the lowest contribution rate for 28 out of 53 flood events, or 53%. The maximum 72-h and 24-h flood volume had the lowest contribution rates for 14 and 10 flood events, respectively.

A significant point is that the results, as presented in Figures 6 and 7, show that the peak level and the total flood volume are the dominant factors of high flood intensity, although flood duration can become a major factor

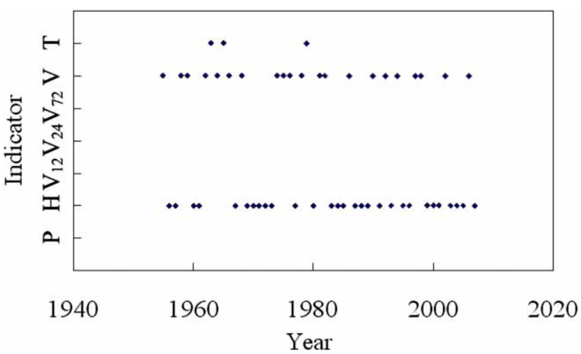

Figure 6 | The indicator corresponding to the maximum contribution rate in the same flood process. P: peak flood discharge; H: peak level; V_{12}: the maximum 12-h volume; V_{24}: the maximum 24-h volume; V_{72}: the maximum 72-h volume; V: the total flood volume; T: flood duration.

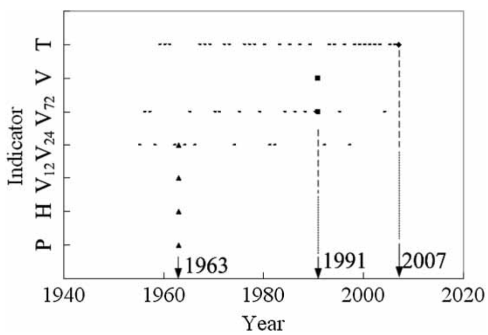

Figure 7 | The indicator corresponding to the minimum contribution rate in the same flood process. ◆ ■ ▲ represents the zero values of the contribution rate. P: peak flood discharge; H: peak level; V_{12}: the maximum 12-h volume; V_{24}: the maximum 24-h volume; V_{72}: the maximum 72-h volume; V: the total flood volume; T: flood duration.

for some flood events. The flood duration apparently had the lowest contribution for flood intensity for most of the historic flood events, although it was a major factor for 6% of the flood events. Wujiang River is a typical mountainous river with a rapid flow, large hydraulic gradient, and a quick rise and fall of flood events. This study showed that the peak flood stages and the total flood volume are the two major factors in the decision-making process for planning and management of the regional hydraulic resources.

The maximum 24-h and 72-h flood volume contributed the least at 26% and 19% of flood events, respectively, and did not occur as a major contributor to the flood intensity; thus, these volumes may not need to be considered for flood management in the Wujiang River.

Assessment of the key factors

Table 1 lists the statistical median, maximum, and minimum of the seven factors used in the analysis. For the 53 flood events, the contribution rates of the peak flood stages ranged from 0% in 1963 to 34.29% in 1991, and the contribution rates of the corresponding peak flood discharges ranged from 0% in 1963 to 21.28% in 1977. The minimum and maximum contribution rate values of flood duration are 0% in 2007 and 72.26% in 1963, respectively. The contribution rate of the total flood volume varied between 0% in 1991 and 40.70% in 1997. The minimum and maximum contribution rates from the maximum 12-h volume are 0% in 1963 and 17.75% in 1993, respectively. Results presented in Table 1 clearly showed that the flood intensity needs to be quantified by multiple factors.

CONCLUSIONS

Quantifying the factors that would significantly contribute to flood intensity is crucial for flood protection and management. Variables involved in flood events are hardly independent. The analysis results might be slightly different if the number of variables chosen for the analysis were different. Nontheless, it is expected that the relative significance in terms of contribution rates for the selected variables can be quantified. This study was based on the commonly used peak flood discharge, the peak flood stage, total flood volume, and four other variables including the flood duration, and 12-h, 24-h, and 72-h flood volumes.

This study has successfully used the projection pursuit model to estimate contribution rates of the seven factors that would have affected flood intensity in the Wujiang River by using 53 years of hydrologic data from the Lishi station. The application of the projection pursuit model showed that the 2006 flood was the largest flood ever recorded in the Wujiang River. Study results also showed that indicator contribution rates for the same indicator were different in different flood events. For the 53 flood events in the Wujiang River, the contribution rate of peak flood stage ranged from 0% in 1963 to 34.29% in 1991, and contribution rates of their corresponding peak flood discharges ranged from 0% in 1963 to 21.28% in 1977.

Table 1 | The result of contribution rate for different indicators from 1955 to 2007 in Wujiang River

Indicator	H	V	V_{12}	P	T	V_{24}	V_{72}
Ave_Cr(%)	22.24	22.18	13.74	13.16	10.11	9.54	9.03
δ	23.40	20.67	14.26	13.45	8.28	9.81	9.15
Year of median value	1985	2007	1996	1996	2003	1975	1987
x_{min}	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Year of minimum value	1963	1991	1963	1963	2007	1963	1991
x_{max}	34.29	40.70	17.75	21.28	72.26	12.80	13.81
Year of maximum value	1991	1997	1973	1977	1963	1993	2007

Ave_Cr: the mean contribution rate. P: peak discharge; H: peak level; V_{12}: the maximum 12-h volume; V_{24}: the maximum 24-h volume; V_{72}: the maximum 72-h volume; V: the total flood volume; T: the flood duration.
The ranks are given in descending order by the mean contribution rate. δ = median; x_{max} = maximum value; x_{min} = minimum value.

This study has shown that peak flood stage and the total flood volume are the dominant factors affecting high flood intensity. Although flood duration can become a major factor for some flood events, it contributed the least to flood intensity for most of the historic flood events. Wujiang River is a typical mountainous river with a rapid flow, large hydraulic gradient, and quick rise and fall of flood events. The Le Chang Valley Hydro-junction project, with a storage capacity of 3.74×10^9 m^3 in 2011, was constructed in the territory of Shaoguan City. In order to better control the flood intensity and risk, decisions are based not only on peak levels, but also on the total flood volume in the Wujiang River. The results are expected to attract more attention from the Le Chang Valley Hydro-junction project regarding flood intensity control and flood risk management.

The maximum 24-h and 72-h flood volumes contributed the least to 26% and 19% of the flood events, respectively, and never occurred as a major contributor to flood intensity; thus, these factors may not even need to be considered for flood management in the Wujiang River. It is worth noting that although the peak flood discharges had good correlation with peak flood stages, the contribution rates for all 53 flood events were much smaller than the peak flood stages had shown. Even though reasons for the low contribution rates of peak flood discharges remain to be explored in future studies, this study showed that peak flood stages and the total flood volume would be the two major factors for the decision-making process in managing regional hydraulic resources.

ACKNOWLEDGEMENTS

This research is financially supported by the National Natural Science Foundation of China (Grant nos. 41501021, 51210013, 51479216), the Public Welfare Project of Ministry of Water Resources (Grant nos. 201201094, 201301002-02), and the project from Guangdong Science and Technology Department (Grant no. 2010B050300010), the project from Guangdong Water Resources Department (Grant no. 2009-39), Key cultivation projects in key reform areas of South China Normal University. Thanks go to Mrs Lisa Sheppard of the Prairie Research Institute at the University of Illinois at Urbana-Champaign for carefully editing the manuscript. The authors would like also to express sincere thanks to the Editor-in-Chief, Professor Chong-Yu Xu, and the anonymous reviewers for their advice for the improvement of this paper.

REFERENCES

Adamowski, K. 2000 Regional analysis of annual maximum and partial duration flood data by nonparametric and l-moment methods. *Journal of Hydrology* **229** (3–4), 219–231.

Ahmed, A. U. & Mirza, M. M. Q. 2000 Review of causes and dimensions of floods with particular reference to Flood 1998: national perspectives. In: *Perspectives on Flood 1998* (Q. K. Ahmad, A. K. A. Chowdhury, S. H. Iman & M. Sarker, eds). University Press Limited, Dhaka, Bangladesh, pp. 67–84.

Amrhein, M., Srinivasan, B., Bonvin, D. & Schumacher, M. M. 1996 On the rank deficiency and rank augmentation of the spectral measurement matrix. *Chemometrics and Intelligent Laboratory Systems* **33** (1), 17–33.

Bai, P., Liu, X., Liang, K. & Liu, C. 2015 Investigation of changes in the annual maximum flood in the Yellow River basin, China. *Quaternary International* **392**, 168–177.

Balke, N. S., Ma, J. & Wohar, M. E. 2013 The contribution of economic fundamentals to movements in exchange rates. *Journal of International Economics* **90** (1), 1–16.

Bhaskar, N. R. & O'Connor, C. A. 1989 Comparison of method of residuals and cluster analysis for flood regionalization. *Journal of Water Resources Planning & Management* **115** (6), 793–808.

Browne, M. J. & Hoyt, R. E. 2000 The demand for flood insurance: empirical evidence. *Journal of Risk Uncertainty* **20** (3), 291–306.

Carolan, M. S. 2007 One step forward, two steps back: flood management policy in the United States. *Environmental Politics* **16** (1), 36–51.

Changnon, S. A., Kunkel, K. E. & Andsager, K. 2001 Causes for record high flood losses in the central United States. *Water International* **26** (2), 223–230.

Cheng, W. S., Ji, C. M. & Liu, D. 2009 Genetic algorithm-based fuzzy cluster analysis for flood hydrographs. In: *Intelligent Systems and Applications, 2009. ISA 2009. International Workshop on. IEEE*, Wuhan, China, pp. 1–4.

Chi, Z. & Dong, S. 2009 A new water quality assessment model based on projection pursuit technique. *Journal of Environmental Sciences (China)* **21** (S1), S154–S157.

Domingo, C., Arcis, R. W., Osorio, E., Toledano, M. & Saurin, J. 2000 Principal component analysis and cluster analysis for the characterization of dental composites. *Analyst* **125** (11), 2044–2048.

Dong, Q. J., Wang, X. J., Ai, X. S. & Zhang, Y. M. 2007 Study on flood classification based on project pursuit and particle swarm optimization algorithm. *Journal of China Hydrology* **27** (04), 10–14.

Droegemeier, K. K., Smith, J. D., Businger, S., Doswell III, C., Doyle, J., Duffy, C., Foufoula-Georgiou, E., Graziano, T., James, L. D., Krajewski, V., LeMone, M., Lettenmaier, D., Mass, C., Pielke Sr, R., Rutledge, S., Ray, P., Schaake, J. & Zipser, E. 2000 Hydrological aspects of weather prediction and flood warnings: report of the ninth prospectus development team of the US weather research program. *Bulletin of the American Meteorological Society* **88** (11), 2665–2680.

Durocher, M., Chebana, F. & Ouarda, T. B. M. J. 2016 Delineation of homogenous regions using hydrological variables predicted by projection pursuit regression. *Hydrology & Earth System Sciences Discussions* **20** (12), 1–23.

French, R. H. & Miller, J. J. 2011 *Flood Hazard Identification and Mitigation in Semi- and Arid Environments*. World Scientific Publishing Company, Singapore.

Friedman, J. H. & Tukey, J. W. 1974 A projection pursuit algorithm for exploratory data analysis. *IEEE Transactions on Computers* **23** (9), 881–890.

Grimalt, J. O., Olive, J. & Gomez-Bellinchon, J. I. 1990 Assessment of organic source contributions in coastal waters by principal component and factor analysis of the dissolved and particulate hydrocarbon and fatty acid contents. *International Journal of Environmental Analytical Chemistry* **38** (2), 305–320.

Hu, T., Wu, F. & Zhang, X. 2007 Rainfall–runoff modeling using principal component analysis and neural network. *Water Policy* **38** (3), 235–248.

Huang, W. & Zhang, X. N. 2011 Projection pursuit flood disaster classification assessment method based on multi-swarm cooperative particle swarm optimization. *Journal of Water Resource & Protection* **3** (6), 415–420.

Javelle, P., Ouarda, T. B. M. J. & Bobee, B. 2003 Spring flood analysis using the flood-duration-frequency approach: application to the provinces of Quebec and Ontario, Canada. *Hydrological Processes* **17**, 3717–3736.

Jie, M. X., Chen, H., Xu, C. Y., Zeng, Q. & Tao, X. E. 2015 A comparative study of different objective functions to improve the flood forecasting accuracy. *Hydrology Research* **47** (4), 718–735.

Jin, J. L., Hong, T. Q. & Li, R. Z. 2005 Objective weighted model for evaluating dynamic eco-environmental comprehensive quality of Chaohu lake basin. *Journal of Hydroelectric Engineering* **24** (5), 99–103.

Jumadar, A. S., Pathirana, A., Gersonius, B. & Zevenbergen, C. 2008 Incorporating infiltration modelling in urban flood management. *Hydrology & Earth System Sciences Discussions* **5** (3), 1533–1556.

Kennedy, L. M. & Basu, M. 2000 Application of projection pursuit learning to boundary detection and deblurring in images. *Pattern Recognition* **33** (12), 2019–2031.

Kharmouz, M. E., Sbaa, M. & Saadi, S. 2013 The contribution of the GIS and the principal component analysis in the study of the cement's dusts impact on the top soils in the central part of the Oujda-Taourirt Corridor (Eastern Morocco). *Open Journal of Air Pollution* **2** (4), 100–106.

Kisss, I., Wessely, E. & Blišťanová, M. 2014 Contribution to logistics of catastrophes in consequence of floods. *Procedia Engineering* **69** (1), 1475–1480.

Latt, Z. Z., Wittenberg, H. & Urban, B. 2015 Clustering hydrological homogeneous regions and neural network based index flood estimation for ungauged catchments: an example of the Chindwin River in Myanmar. *Water Resources Management* **29**, 1–16.

Liu, D., Zhao, D., Liang, X. & Wu, Q. 2016 Research on evaluating water resource resilience based on projection pursuit classification model. *Applied Water Science* **6** (1), 97–105.

Mehdi, N., Ghodrat, S., Mehdi, S. & Adel, N. T. 2012 Artificial fish swarm algorithm: a survey of the state-of-the-art, hybridization, combinatorial and indicative applications. *Artificial Intelligence Review* **42** (4), 965–997.

Miloudi, N., Hammoum, H., Bouzelha, K., Metna, B. & Pantet, A. 2017 Contribution to assess the water storage tank vulnerability by principal component analysis. *Journal of Civil Structural Health Monitoring* **7**, 275–289.

Noori, R., Khakpour, A., Omidvar, B. & Farokhnia, A. 2010 Comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic. *Expert Systems with Applications* **37** (8), 5856–5862.

Noori, R., Karbassi, A. R., Moghaddamnia, A., Han, D., Zokaei-Ashtiani, M. H., Farokhnia, A. & Gousheh, M. G. 2011 Assessment of input variables determination on the SVM model performance using PCA, Gamma test, and forward selection techniques for monthly stream flow prediction. *Journal of Hydrology* **401** (3–4), 177–189.

O'Connor, J. E. & Costa, J. E. 2003 *Large Floods in the United States: Where They Happen and why. US Geological Survey Circular 1245*. USGS, Reston, VA, USA.

O'Connor, J. E. & Costa, J. E. 2004 *The World's Largest Flood, Past and Present: Their Causes and Magnitudes. US Geological Survey Circular 1254*. USGS, Reston, VA, USA.

Qin, B. Y. & Lin, X. K. 2012 Construction of response surface based on projection pursuit regression and genetic algorithm. *Physics Procedia* **33**, 1732–1740.

Rajeevan, M., Pai, D. S., Kumar, R. A. & Lal, B. 2007 New statistical models for long-range forecasting of southwest monsoon rainfall over India. *Climate Dynamics* **28** (7), 813–828.

Reza, A. & Bojnourd, I. 2014 Empirical study of artificial fish swarm algorithm. *International Journal of Computing, Communications and Networking* **3** (1), 1–7.

Shabri, K., Mamta, G., Rohit, K., Rohan, S., Dalam, S., Akash, P. & Dileshwar, K. S. 2014 Estimation of maximum discharge for small catchment area. *International Journal of Computer Science and Network* May, 2277–5420.

Singh, R. B. & Singh, S. 2011 Rapid urbanization and induced flood risk in Noida, India. *Asian Geographer* **28** (2), 147–169.

Sun, H., Guo, H. X., Hu, J. L. & Zhu, K. J. 2012 Comparative study on economic contribution rate of education of China. *Applied Soft Computing* **12**, 2106–2113.

Swades, P. & Surajit, L. 2012 Flood intensity and potential flood loss estimation in Dwarka River basin of eastern India. *International Journal of Geology, Earth and Environmental Sciences* **2** (1), 116–122.

Tomašić, I., Trobec, R. & Magjarević, R. 2014 Principal component analysis of differential electrocardiographic leads – a contribution to the synthesis of ECGs. In: *The International Conference on Health Informatics*, pp. 260–263.

Tu, X. Y. 1999 Artificial animals for computer animation. *Lecture Notes in Computer Science* **1635**, 361–364.

Wang, S. J. & Zhang, X. L. 2010 Projection pursuit dynamic cluster model and its application to water resources carrying capacity evaluation. *Journal of Water Resource & Protection* **2** (5), 449–454.

Wang, S. J., Yang, Z. F. & Ding, J. 2004 Projection pursuit method of comprehensive evaluation on groundwater resources carrying capacity in Guanzhong plain. *Resources Science* **26** (6), 104–110 (in Chinese).

Wang, L. N., Chen, X. H., Shao, Q. X. & Li, Y. 2015 Flood indicators and their clustering features in Wujiang River, South China. *Ecological Engineering* **76**, 66–74.

Ward, S. M. & Paulus, G. 2013 Augmenting Austrian flood management practices through geospatial predictive analytics: a study in Carinthia. *Natural Hazards & Earth System Science* **13** (6), 1445–1455.

Yazdani, D., Toosi, A. N. & Meybodi, M. R. 2010 Fuzzy Adaptive Artificial Fish Swarm Algorithm. In: *AI 2010: Advances in Artificial Intelligence, Australasian Joint Conference*, Adelaide, Australia, 6464, pp. 334–343.

Yazdani, D., Akbarzadeh-Totonchi, M. R., Nasiri, B. & Meybodi, M. R. 2012 A new artificial fish swarm algorithm for dynamic optimization problems. In: *WCCI 2012 IEEE World Congress on Computational Intelligence*, June 10–15, Brisbane, Australia, pp. 472–479.

Zhao, J., Jin, J., Guo, Q., Liu, L., Chen, Y. & Pan, M. 2014 Dynamic risk assessment model for flood disaster on a projection pursuit cluster and its application. *Stochastic Environmental Research and Risk Assessment* **28** (8), 2175–2183.

Zheng, G. R. & Lin, Z. C. 2012 A winner determination algorithm for combinatorial auctions based on hybrid artificial fish swarm algorithm. *Physics Procedia* **25**, 1666–1670.

First received 3 March 2017; accepted in revised form 18 July 2017. Available online 13 September 2017