Copula Based Rainfall Designing for Spatial Flood Risk Assessment

Xinyu JIANG[1], Hirokazu TATANO[2]

[1] Researcher, Disaster Prevention Research Institute, Kyoto University
(Gokasho,Uji,Kyoto611-0011, Japan)
E-mail: jiang.xinyu.4u@kyoto-u.ac.jp
[2]Member of JSCE, Professor, Disaster Prevention Research Institute, Kyoto University
(Gokasho,Uji,Kyoto611-0011, Japan)
E-mail: tatano@imdr.dpri.kyoto-u.ac.jp

Spatial distribution of flood risk is important information for integrated urban flood risk management. Conventional flood risk assessment deal with flood risk sources independently. This leads flood risk in the flood plain underestimated. Even in the floodplain which has no risk from sea, flood from river channels and inundation caused by the shortage of drainage capacity should be simultaneously considered. For itegraed flood risk management, it is necessary to establish a methodology to estimate flood risk distribution over floodplain. In this paper, a method of rainfall designing for spatial flood risk assessment which considering joint effects of multiple flood sources is proposed. The concept of critical rainfall duration determined by concentration time of flood is introduced to connect responses characteristics of different flood sources with rainfall. Copula method is then adopted to capture the correlation of rainfall amount within different critical rainfall durations. The rainfall are designed taking advantage of copula structure of correlation and marginal distribution of rainfall amount within different critical rainfall durations. A case study in Otsu river basin, Osaka prefecture, Japan is conducted to demonstrate this methodology.

Key Words : *spatial flood risk assessment, rainfall designing, copula,multiple flood sources*

1. INTRODUCTION

Flood is one of the most serious disasters in the word. Annually, flooding affects about 520 million people and their livelihoods, claiming about 25,000 lives worldwide. The annual cost to the world economy, of flooding and other water-related disasters, is between $50 and $60 billion[1)]. To prevent people from flood disaster and reduce economic loss, integrated flood risk management which include different types of countermeasures, multiple stakeholders and authorities is required[2,3)]. In order to manage flood risk efficiently, it is necessary to assess spatial distribution of flood on whole flood plain[4)]. Spatial flood risk assessment should reflect all risk information derive from multiple flood sources: rivers, drainage, coast etc. which may affect urban areas. Each of the flood sources is studied independently in conventional research, therefore, joint effects of multiple flood sources are ignored which may lead to spatial flood risk underestimated. As the risk is represented by probability distribution of loss[5)], to assess spatial flood risk, a challenge we are facing is how to evaluate the joint probability distribution of maximum water depth which may derive from different flood sources, furthermore the joint probability distribution of loss.

Although floods may comes from many sources, for most area, heavy rainfall could be the main root cause and most flood risk assessment are start from rainfall. Scholars usually use the following procedure to estimate probabilistic flood risk curve: 1. Design rainfall corresponding to certain return period. 2. Put the designed rainfall into integrated or separated rainfall-runoff-inundation model to simulate water depth at each place in the area. 3. Change dyke break or overtopping point to achieve many scenarios of flooding. 4. Calculate losses under each scenarios and calculate risk curve.

As the first step of the flood risk assessment procedure, rainfall connects the flood simulations to statistic risk analysis. There are variety of methods of rainfall designing exist in the literatures. Most of rainfall design methods could be classified into four categories: specification of simple geometrical shapes anchored to a single point of the intensity duration frequency (IDF) curve; use of the entire IDF curve; use of standardized profiles obtained directly from rainfall records, and simulation from stochastic models.

Conventional methods of rainfall designs such as “design from intensity–duration–frequency (IDF) curve” or “extend from typical rainfall” are not proper for spatial flood risk assessment which considering multiple flood sources because two reasons: on one hand, intensity–duration–frequency (IDF) curve is usually determined by means of statistical analysis of univariate of mean intensity of a rainfall. Only one type of rainfall could be designed from these methods under certain return period. It hardly to says this type of rainfall is most likely to occur under this return period. On the other hand, flood which affect the place may comes different sources, for example, larger rivers, smaller rivers or urban drainages. Different flood sources may have different responses to designed rainfall: larger rivers may require a longer duration of rainfall to produce a flood peak while urban drainage may be more sensitive to shorter duration of rainfall or peak rainfall. The rainfall design should consider the relationship between response characteristics of different flood sources.

In this paper, a copula based method of rainfall designing for spatial flood risk assessment which considering joint effects of multiple flood sources is proposed. The concept of critical rainfall duration on the basis of concentration time of flood is introduced to connect responses characteristics of different flood sources with rainfall. Copula method is then adopted to capture the correlation of rainfall amount within different critical rainfall durations. The rainfall are designed taking advantage of correlation and marginal distribution of rainfall amount within different critical rainfall durations. Therefore, the joint effects of multiple flood sources are reflected in the design rainfall.

This paper is arranged in the following way. Section 2 introduces the proposed rainfall design procedure and related concepts and methods. Section 3 presents a case study which apply this rainfall design procedure to the Otsu River basin in Osaka, Japan for spatial flood risk assessment. In section 4, conclusions are presented.

2. METHODOLOGY

The rainfall design in this paper is for spatial flood risk assessment. Design rainfall is an input to hydrological and hydrodynamic model. By using these models, water depth at each place in an area could be calculated. The integrated spatial flood risk assessment focus on risk at each place in an area and consider flood risk from all possible flood sources. The rainfall design should consider the relationship between response characteristics of different flood sources.

(1) Concentration time of flood

The response of a catchment to a rainfall event can be measured by concentration time of flood which defined as the time required for disturbance of rainwater to propagate from the top of slope at the remotest portion of the basin in the sense of dynamics to the outlet[6]. It is reasonable to set concentration time of flood as critical rainfall duration, in which rainfall include peak rainfall will form flood peak volume. Many previous studies on calculating concentration time of flood have been reported such as Kraven formula, uniform flow velocity formula, Public Works Research Institute formula, and Kadoya formula[7]. Concentration time of flood could be different between different flood sources, critical rainfall duration could be different. Therefore, to assess spatial flood risk considering joint effect of multiple flood sources, it is significant to understand the correlation of different critical rainfall durations and their joint distributions.

(2) Copula method

Copula method is a popular and flexible way to measure correlation as well as construct multivariate distributions. Copulas are functions that join or “couple” multivariate distribution functions to their one-dimensional marginal distribution functions[8]. In the bivariate case, the joint cumulative distribution function H(x,y) of any pair (x,y) of continuous random variables can be written in the form:

$$H(x,y) = C(F(x),G(y)) \;\; x,y \in \mathrm{R} \tag{1}$$

Where F(x) and G(y) are continuous marginal distributions, so that the C:[0,1]→[0,1] such that for all is copula[9]. This method separate the joint distribution into a copula function and marginal distributions and provide an advantage that the selection of an appropriate model for the dependence between varieties, represented by the copula, can then proceed independently from the choice of marginal distributions. As for the basic theory and concepts of copula, readers may refer to the monographs by Joe and Nelsen for detail[8,10]. For construction of high dimensional copula, such as nested Archimedean construction (NAC) and pair copula construction (PCC), readers may refer to Aas and Berg, Savu and

Trede and Czado[11,12,13].

In this paper, the copula method is adopted to analyze the correlation of rainfalls within critical rainfall durations of different flood sources and to built the joint distributions of rainfall amount within critical rainfall durations.

(3) Methodology framework

To sum up, the rainfall design procedure for spatial flood risk assessment which considering joint-effects of multiple flood sources is formed, as is shown in Fig.1. Firstly, the rainfall data and basin information should be collected. Based on basin information, the flood assessment area could be defined. Note that the basin refers to the whole area include mountain runoff area and flood assessment area. From the view point of integrated flood risk assessment, the whole process rainfall runoff inundation is important. Centered consider the risk assessment area, the flood risk sources could be traced and the concentration time of flood from flood risk sources to flood risk assessment area could be calculated. As is stated above, concentration time of flood could be thought as critical rainfall duration, in which rainfall include peak rainfall will form flood peak volume. On the other hand, rainfall series are proceed. The basin average rainfall should be calculated because many flood risk assessment at small scale require basin average rainfall as input. It could be achieved by interpolation methods such as, Thiessen polygons methods, Inverse Distance Weighted method, polynomial methods or Kriging methods. Then, the rainfall time series would be divided into rainfall events through interval time method or mean-time method. Up to this, the rainfall data is prepared for statistical analysis. Next, the correlation of rainfall within different critical rainfall durations will be estimated through copula method, for example, 1h rainfall, 2h rainfall, 3h rainfall etc. For each duration of rainfall, the probability distribution will be fit. The joint distribution of rainfall within different durations could be constructed by the copula correlation structure and marginal distributions. The joint probability distribution will be used to generate critical rainfall according to return periods. Finally, the generated correlated critical rainfall is used to form rainfall events. The whole procedure will be realized in a case study in next section.

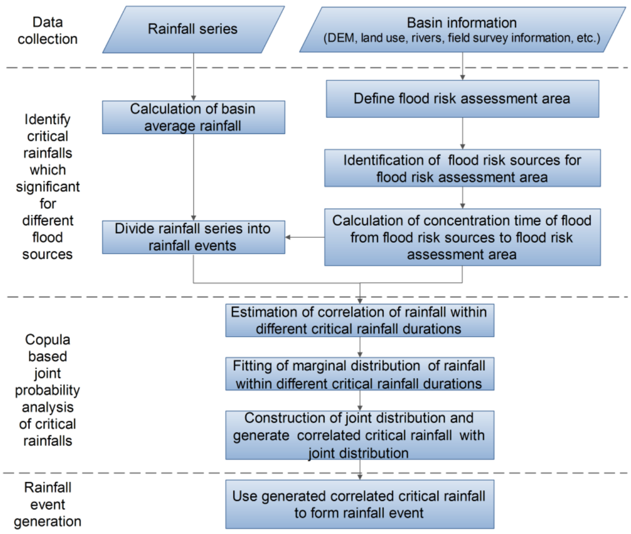

Fig.1 Rainfall design procedure for spatial flood risk assessment which considering joint-effects of multiple flood sources.

3. CASE STUDY

(1) Study area

The study area is located in the southern part of Osaka prefecture, Japan. Otsu River is a river of Osaka prefecture rising in the Katuragi Mountain and flowing about 68 km westward to the Osaka bay. It consist of 5 branches include Chichioni river, Higashimakio river, Makio river, Matsuo river and Ushitaki river. It is the largest secondary drainage in Osaka prefecture. The upstream of the basin is mountain area and covered by natural landscape; the middle part of the basin is hilly area and partly developed; the downstream of the basin is well developed urban area. The sketch map of Otsu River basin is shown in Fig.2. In this research, the proposed method is used to design a average basin rainfall for Otsu river basin for the purpose of spatial flood risk assessment. Geographic data include digital map of Osaka prefecture, digital elevation data (DEM) and river system are collected from Geospatial Information Authority of Japan. The rainfall data include four rain gauging stations with 49 years hourly rainfall records and historical flood disaster record and field survey data are provide by River management department of Osaka prefecture.

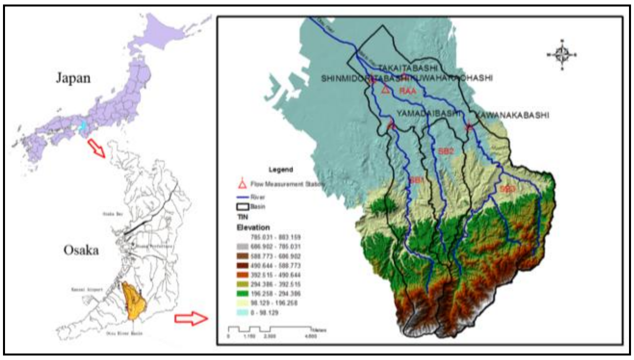

Fig.2 Sketch map of study area.

(2) Identify critical rainfalls

Historical record shows that this area prone to flood disaster. From 1950 to 2011, 14 flood disaster events were recorded. Local inundation and river flood is the main cause of flood disaster. Especially, the downstream of the basin is flat, three river flows are concentrated, which gives it a high probability of flooding by multiple rivers and local inundation. In addition, the population and properties are concentrated in the downstream of basin. It is meaningful to consider this area first as a risk assessment area. For the risk assessment area, flood sources include river flooding from Ushitaki River, which is controlled by runoff in the upper part of Ushitaki sub-basin (SB1); river flooding from Matsuo River, which is controlled by runoff in the upper part of Matsuo sub-basin (SB2); river flooding from Makio River, which is controlled by runoff in upper part of Makio sub-basin (SB3); and local inundation from urban drainage or slope flow, which controlled by runoff in the flood risk assessment area (RAA).

According to empirical Kraven formula, the flood concentration time of Ushitaki Basin is 2h, that of Matsuo Basin is 1.6h, and that of Makio Basin is 2.7h. Because there is no big reservoir or dam in the runoff areas, the flood concentration time may imply that 2h rainfall will be critical for Ushitaki and Matsuo basins to produce flood peak to risk assessment area, and 3h rainfall could represent that in Makio Basin to produce flood peak to risk assessment area. While for the flood concentration time in risk assessment area, 1h rainfall is considered to be critical. Thus, analysis of joint probability of flooding from multiple sources become analysis of joint probability rainfall of 1h, 2h, and 3h rainfall under the assumption of basin average rainfall.

Three hour is taken as interval to divide basin average rainfall series into rainfall events. Although all of rainfall data could be used to evaluate rainfall dependence structure, the flood risk analysis concern more on some extreme rainfalls. Annual maximum rainfall events corresponding to 1h, 2h and 3h duration are selected.

(3) Copula based joint probability analysis of critical rainfall

In case study, 1h, 2h and 3h rainfall are considered, therefore, three dimensional copula analysis are required. There are two types of method to construct high dimensional copula: nested Archimedean construction and pair copula construction. Compare with nested Archimedean construction, pair copula construction is a more flexible method for multivariate copula because it adopts a hierarchical idea and takes advantage of density function. In our research, the peak rainfall could be the key variable, the relationship between one hour rainfall and two hours rainfall or the relationship between one hour rainfall and three hours rainfall is relative important than the relationship between two hour rainfall and three hours rainfall.

Some 20 types of copulas introduced by Genest and Favre which may be suitable for hydrological study are considered, include Archimedean copulas, extreme-value copulas, meta-elliptical copulas and other miscellaneous families of copulas[14]. For the selection of Copula, the Akaike information criterion (AIC) is adopted. The AIC shows that survival Gumbel with parameter 3.357, Gaussian copula with parameter 0.804 and BB7 copula with parameter 2.923 and 3.451 can properly fit 1h-2h rainfall correlation, 1h-3h rainfall correlation, and conditional 2h-3h rainfall correlation respectively. Fig.3 shows the 3d scatter points of pseudo data and fitted copula density of 1h-2h rainfall, 1h-3h rainfall and conditional 2h-3h rainfall, from which the correlation of rainfall in different durations could be illustrated.

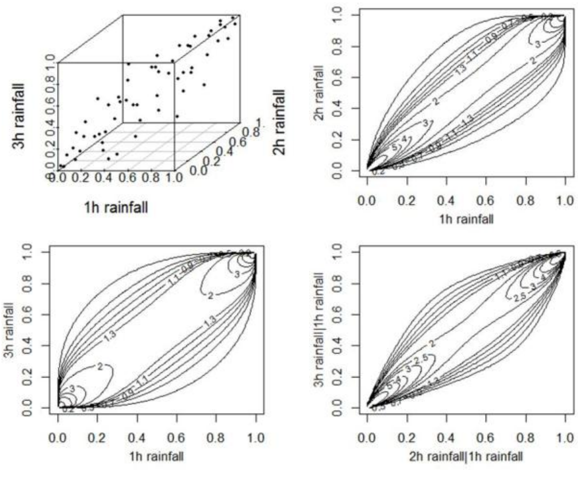

Fig.3 The 3d scatter points of pseudo data and fitted copula density of 1h-2h rainfall, 1h-3h rainfall and conditional 2h-3h rainfall.

The next is estimation of marginal distributions. Many previous studies on fitting of extreme rainfall distribution have been done and several types of distribution are found well fit rainfall data. However, no distribution is universally fitted to all rainfall data due to various nature of rainfall, different purposes of study, different locations, etc. In this study, a set of distributions include Pearson distribution families, Generalized Pareto distribution (GP), Generalized Extreme Value distribution (GEV), Exponential distribution (EXP), Gamma distribution (GM), log-normal distribution, and Weibull distribution families are selected as candidates and tested by Kolmogorov–Smirnov test and as well as AIC. The tests indicate that for annual maximum1h rainfall, the best fit distribution is Lognormal with parameter (3.132, 0.337); for annual maximum 2h rainfall, the best fit distribution is Pearson 3 with parameter (2.155, 16.406, 9.109); for annual maximum 3h, the best fit distribution is Lognormal with parameter (3.697, 0.362).

After estimate correlation structure and marginal distribution, the joint distribution could be constructed. According to Sklar's theorem, the joint distribution could be separated into copula model with marginal distribution as is shown in fumula (2).

$$F(x_1,x_2,x_3)=C_{1,2,3}(F_1(x_1),F_2(x_2),F_3(x_3)) \quad (2)$$

Where, in our study, x_1, x_2, x_3 donates 1h rainfall, 2h rainfall and 3h rainfall. Using the chain rule, we have

$$f(x_1,x_2,x_n)=c_{1,2,3}(F_1(x_1),F_2(x_2),F_3(x_3)) \cdot f_1(x_1)\cdot f_2(x_2)\cdot f_3(x_3) \quad (3)$$

where $c_{1,2,3}$ denotes the densities of $C_{1,2,3}$. The copula densities $c_{1,2,3}$ could be calculated through PCC as is shown in formula (4).

$$c_{1,2,3}(x_1,x_2,x_3)=c_{1,2}(F_1(x_1),F_2(x_2)) \cdot c_{1,3}(F_1(x_1),F_3(x_3)) \cdot c_{2,3|1}(F(x_2\,|\,x_1),F(x_3\,|\,x_1)) \quad (4)$$

Therefore, once marginal distributions and copula model are determined, the joint distribution could be constructed. The algorithms of sampling value from vine copulas are proposed by Aas et al.[15]. Take advantage of these algorithms, random copula value could be generated. Then based on the marginal distributions, the random value of the critical rainfall could be obtained. Fig.4 shows 10000 random values generated by copula functions and the observed real rainfall data. From this figure, it is can be found that the correlation of real rainfall data are captured by copula model. It is reasonable to adopt the simulated random value to flood risk assessment.

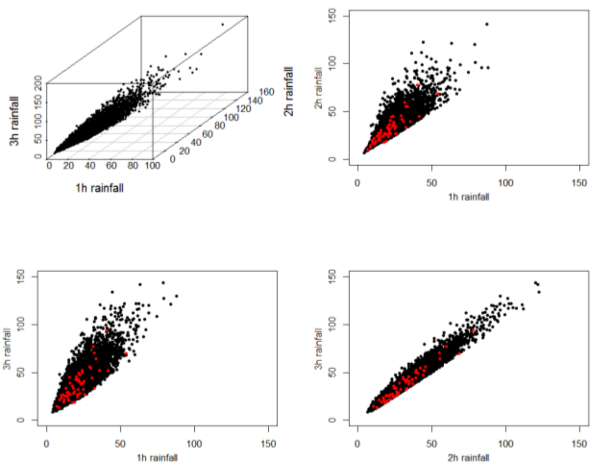

Fig.4 10000 random rainfall values generated by copula model. The black points are random value; the red points are observed real rainfall data.

The flood risk assessment requires information of probability distribution of water depth which simulated from corresponding rainfall event. The simulated random rainfall points include 4 dimensional information of joint probability, amount of 1h, 2h, and 3h rainfall. To complete the rainfall events, we assume critical appear at the center part of rainfall event. Then, the amount of 1h, 2h, and 3h rainfall could be located, as is shown in left of Fig.5. Since the generated points also have information of joint probability, the designed rainfall will share the same probability. As for the left part of rainfall event, although they will also affect runoff, compare with critical rainfall duration, they may make less contribution to flood peak. Therefore, the left part could be simple completed by statistical average of historical rainfall events.

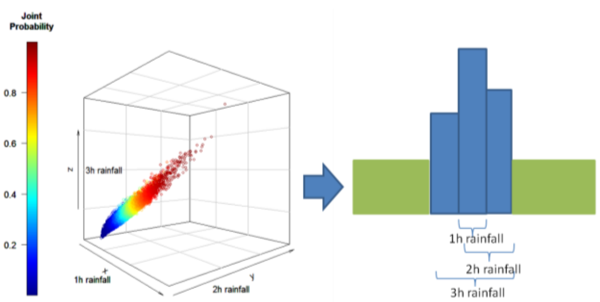

Fig.5 Generation of rainfall events from simulated correlated critical rainfall.

Taking advantage of this rainfall designing method, the basin average rainfall are designed for case study area. Because the joint probability is adopted, the rainfall values of certain return period could be a surface, more than one rainfall value could be expected at certain return period. In Fig.6, 5 rainfall events for 20, 50 and 100 years return periods are generated respectively.

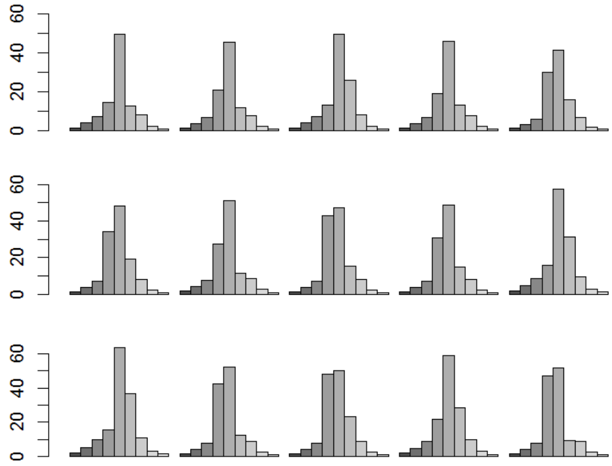

Fig.6 5 rainfall cases for 20 years (top), 50 years (middle) and 100 years (bottom) return periods respectively.

(4) Apply generated rainfall event spatial to flood risk assessment

The generated rainfall events are used to drive cascading flood simulations to assess flood risk at each palce of floodplain. A GIS-based integrated rainfall–runoff–inundation model was developed. The model applied unstructured irregular meshes and simplified 2D shallow water equations to flood and inundation simulation in the risk assessment area. Moreover, it applied hydrological analysis and 1D kinematic wave equations to rainfall–runoff simulation in the runoff area and was able to simulate runoff, flooding and inundation together. GIS is fully adopted in data management and data processing as well as result visualization. The simulation model set up and validation is shown in Fig.7.

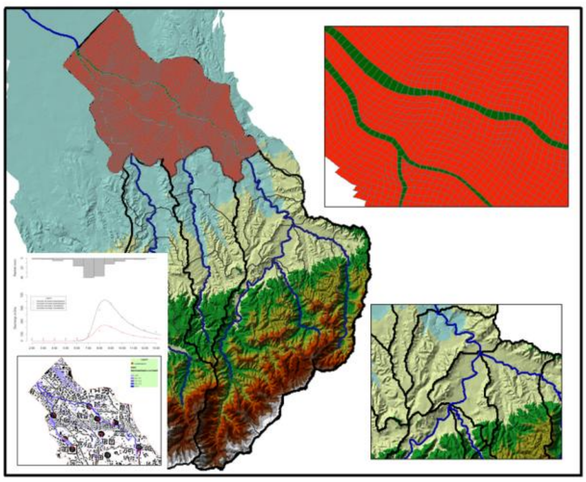

Fig.7 GIS-based integrated rainfall–runoff–inundation model set up and validation.

Simulating each rainfall event, inundation could be estimated and further loss could be calculated with the addition information of fragirity curve and exposure distribution. Because the joint probability is adopted, for each return period, more than one rainfall events could be generated. This characteristic of the methodology enable us estimate not only event curve but also risk curve considering uncertainty. Taking mesh No. 5548 for example, the loss and return period of events were plotted as shown in Fig.8, and an event curve was drawn using the average loss of each return period and risk curve considering uncertainty are also shown.

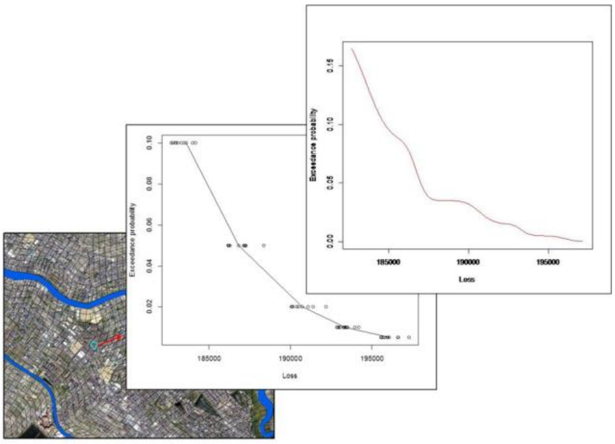

Fig.8 Event curve and risk curve of mesh No. 5548.

For each mesh, risk curve could be calculated. To map spatial distribution of flood risk, the expected loss calculated from the risk curve of each mesh. Fig.9 shows spatial distribution of flood risk in terms of expected loss.

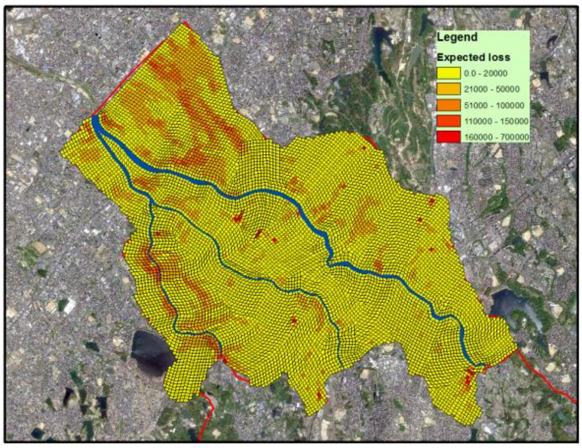

Fig.9 Map of spatial distribution of flood risk in terms of expected loss

4. CONCLUSION

This paper proposed a copula based methodology for rainfall designing for spatial flood risk assessment considering multiple flood risk sources. A case study in Otsu river basin, Osaka prefecture, Japan is conducted to demonstrate this methodology.

Copula method could be a proper method to build joint probability distribution for rainfall designning. Apply copula constructing joint probability to rainfall analysis shows an advantage that even the same return period, different rainfall events could occurs. For example, in the case study, it is clearly seen that even joint probability is the same, marginal probability could be quite different, and rainfall events could be different. It also shows the necessity to consider the joint probability rather than a single marginal distribution. In the case study, the joint effects from multiple risk sources are represented by different duration of rainfall amount. The joint probability offers more cases of rainfall event which is quite important for flood risk assessment.

ACKNOWLEDGMENT: This work was supported under the framework of the 'Precise Impact Assessments on Climate Chang' of the Program for Risk Information on Climate Change (SOUSEI Program) supported by the Ministry of Education, Culture, Sports, Science, and Technology-Japan (MEXT); It was also supported by the program of 'Study on Spatial Flood Risk Assessment Considering spatial-temporal correlation' supported by Japan Institute of Country-ology and Engineering (JICE).

REFERENCES

1) Teegavarapu R S V. Floods in a Changing Climate: Extreme Precipitation. *Cambridge University Press*, 2012.
2) Tatano H. Major characteristics of disaster risk and its management strategies. *Sociotechnica*, 1: 141-148, 2003.
3) Okada N. Urban diagnosis and integrated disaster risk management. *Journal of Natural Disaster Science*, 26(2): 49-54, 2004.
4) Taki K, Matsuda T, Ukai E, et al. Method for evaluating flood disaster reduction measures in alluvial plains[J]. *Journal of Flood Risk Management*, 6(3): 210-218, 2013.
5) Knight F H. Risk, uncertainty and profit. *Courier Dover Publications*, 2012.
6) Kadoya, M, Fukushima, A. Concentration Time of Flood In Small or Medium River Basin. *Disaster Prevention Research Institute Annuals*, 19: 143-152, 1976.
7) Japan Society of Civil Engineers. The Collection of Hydraulic Formulae, 1986. (In Japanese)
8) Nelsen, R. An Introduction to Copulas, 2nd edition. *Springer, New York*, 2006.
9) Sklar A. Fonctions de repartition an dimensions etleurs-marges. *Publ Inst Statist Univ Paris* 8:229-231, 1959.
10) Joe H. Multivariate Models and Dependence Concepts. *Chapman &Hall, London*, 1997.
11) Aas K, Berg D. Models for construction of multivariate dependence—a comparison study . *The European Journal of Finance*, 15(7-8): 639-659, 2009
12) Savu C, Trede M. Hierarchies of Archimedean copulas . *Quantitative Finance*, 10(3): 295-304, 2010.
13) Czado C. Pair-copula constructions of multivariate copulas /Copula theory and its applications. *Springer Berlin Heidelberg*: 93-109, 2010.
14) Genest C, Favre A C. Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask. *J Hydrol Eng.* 12 (4): 347-366, 2007.
15) Aas K, Czado C, Frigessi A, et al. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44(2): 182-198, 2009.