

Verification of ArcGIS for flood hazard mapping: A case study of Cholburi Province, Thailand

Supatchaya Chuanpongpanich
Faculty of Engineering at Sriracha,
Kasetsart University Sriracha Campus
Cholburi, Thailand
Email: supatchaya@eng.src.ku.ac.th

Tawatchai Tingsanchali*
Faculty of Engineering at Sriracha,
Kasetsart University, Sriracha Campus
Cholburi, Thailand
Email: tawatchai2593@gmail.com

Abstract—Flood hazard maps of a study area can be constructed by using ArcGIS. Generally, the weights of input data are assumed without verification. The objective of this study is to develop a novel method to verify the ArcGIS for flood hazard mapping given the input data, such as: rainfall, ground slope, ground terrain elevation, land use and soil types. To calibrate and verify the ArcGIS, the weight of individual input data is first assumed and adjusted systematically and iteratively to determine the best set of weights of all input data for computing the flood hazard maps. The calibration and verification of the weights are based on comparison of the computed flood inundation areas with the satellite observed data. The final weights are obtained when satisfactory agreement between the computed and observed flood hazard maps is obtained.

A case study was carried out for flood hazard mapping in Cholburi Province, Thailand. Five input data to ArcGIS are specified as mentioned above. The weights of all input data are assumed and adjusted systematically in the calibration to obtain satisfactory agreement between the computed and actual flood inundation maps in 2013 and 2010. The same weights obtained from calibration are verified by using them to compute the inundation map in 2008 and compared with the satellite observed inundation. The comparison shows satisfactory agreement on calibration and verification of ArcGIS. It is recommended to apply this method to other floods and in other areas.

Keywords—ArcGIS verification; flood hazard maps; Cholburi, Thailand

I. INTRODUCTION

Cholburi Province in East Thailand has an area of 4,634 km² (Fig. 1). The topography of the province is mountainous in the east, northeast and south of the province where low lying areas are in the northwest and west. After heavy rainfalls, floods occurred and caused large damages in northern part of Cholburi. The development of the flood hazard map for the Cholburi province will be useful for planning of land use and for flood damage mitigation. Flood hazard mapping is a vital component for appropriate land use planning in flood-prone areas. It creates easily-read, rapidly-

accessible charts and maps which facilitate the identification of areas at risk of flooding and also helps prioritize mitigation and response efforts [1]. A previous study [2] has been done to estimate flood hazard area in the Cholburi province based on a matching index computed from ArcGIS for overall inundation area in the 10-year period of 2008-2017. The five types of input data considered in their study were rainfall, ground slope, ground elevation, land use and soil type. But the study [2] did not calibrate and verify the weights of input data. This may lead to error as the weights of input factors control the score of the matching index for inundation area. Other previous studies such as the study by [3] on flood risk area in the upper Pasak river basin in Lomsak, Phetchabun Province and [4] on flood risk area around western sub-basins area of Songkhla Lake in Phatthalung Province using GIS did not calibrate and verify the weights of input factors. Therefore it is in doubt whether the weights of the input factors are correct or not.

II. PURPOSE OF STUDY

The purpose of this study is to calibrate and verify the ArcGIS model by searching the best set of weights of the input factors that work well for flood inundation in the considered calibration and verification years. In this study, the annual flood inundations in 2013 and 2010 were considered for calibration and in 2008 for verification. The successful calibration and verification will assure the validity and reliability of the ArcGIS in estimating flood inundation in the study area in other flood years and elsewhere.

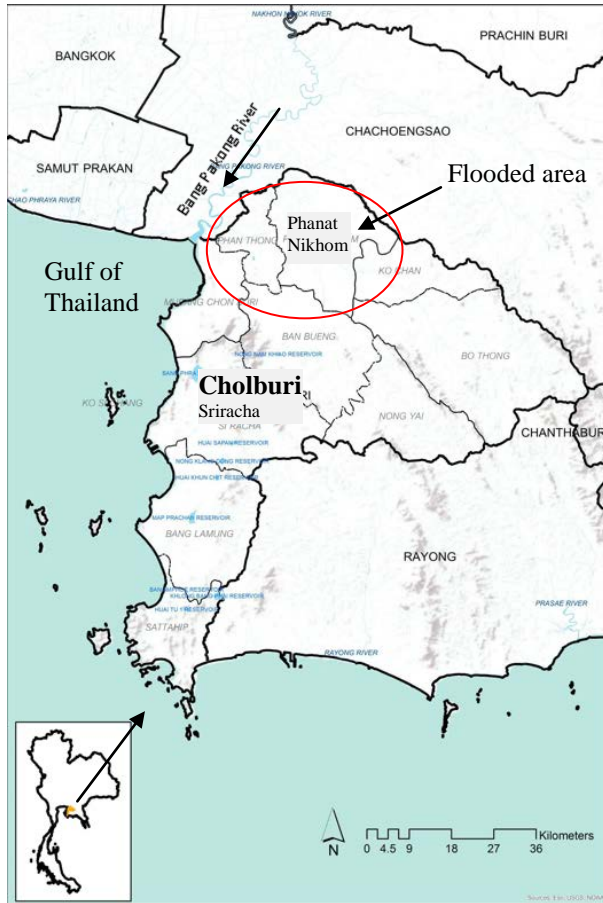


Fig. 1. The flooded area under study, Choburi Province

III. FLOOD HAZARD MAPPING USING ARCGIS

Computing Matching Index

ArcGIS version 10.5 which is a Geographic Information System application for map construction is used to analyze the related parameters and flooded areas by weighting technique [5] and [6]. The matching index M is computed and compared with the satellite observed flood inundation area. Five input factors are used namely: rainfall with assigned index number $i=1$, ground slope with $i=2$, ground elevation with $i=3$, land use with $i=4$ and soil type with $i=5$. Equation 1 is used to compute the matching index M based on point score analysis [7] to represent the magnitude of flood hazard inundation area.

$$M = W_1R_1 + W_2R_2 + W_3R_3 + \dots + W_iR_i + \dots + W_nR_n \quad (1)$$

Where W_i is the weight of i^{th} input factor; R_i is the scale value of the magnitude of the i^{th} input factor; n is the total number of input factors and n is equal to 5. The weight W is set equal to 5 for the most influential input factor. For the least influential input factor, W is set equal to 1.

Each input factor is categorized into 5 classes with intervals according to the ranges of field data as shown in Table 1. Each class of input factor is designated by a scale value R , e.g. $R=1$ for very low significance, 2 for low, 3 for moderate, 4 for high and 5 for very high.

For example, the flood inundation is increased when the amount of rainfall increases. Therefore, the highest scale $R=5$ is given to the rainfall of more than 200 mm. The lowest scale $R=1$ is given to the lowest range of rainfall that is less than 50 mm as shown in Table 1. The scales R of the classes of the other input factors are specified according to field data range and are kept unchanged throughout this study.

TABLE 1. CLASSES AND INTERVALS OF INPUT FACTORS AND ASSIGNED SCALE VALUES

Input factor and index i	Classes of input	Inputs and their intervals	Scale of input factor, R^*
Rainfall, $i=1$	1	Class 1: Less than 50 mm	1
	2	50 – 100 mm	2
	3	100 – 150 mm	3
	4	150 – 200 mm	4
	5	More than 200 mm	5
Ground slope, $i=2$	1	More than 20%	1
	2	15-20%	2
	3	10-15%	3
	4	5-10%	4
	5	Less than 5%	5
Elevation, $i=3$	1	High mountain	1
	2	Mountainous	2
	3	Hilly terrain or plateau	3
	4	Low land	4
	5	Flood plain	5
Land use, $i=4$	1	Forest	1
	2	Industrial	2
	3	Urban	3
	4	Agricultural	4
	5	Rivers or canals	5
Soil type, $i=5$	1	Very well drained	1
	2	Well drained	2
	3	Moderately drained	3
	4	Poorly drained	4
	5	Very poorly drained	5

*scale of 1 means lowest significance to flood inundation, and of 5 means highest significance

Spatial Rainfall Distribution over Study Area

Due to non-uniform spatial distribution of rainfall, annual rainfalls at 21 stations in Choburi Province were interpolated to obtain spatial average rainfall for input to

ArcGIS. This was done by using Inverse Distance Weighted (IDW) method from ArcGIS as shown in Fig. 2. This method is a spatial interpolation technique that can estimate spatial average rainfall over the study area based on the distance from known rainfall values at many stations [8].

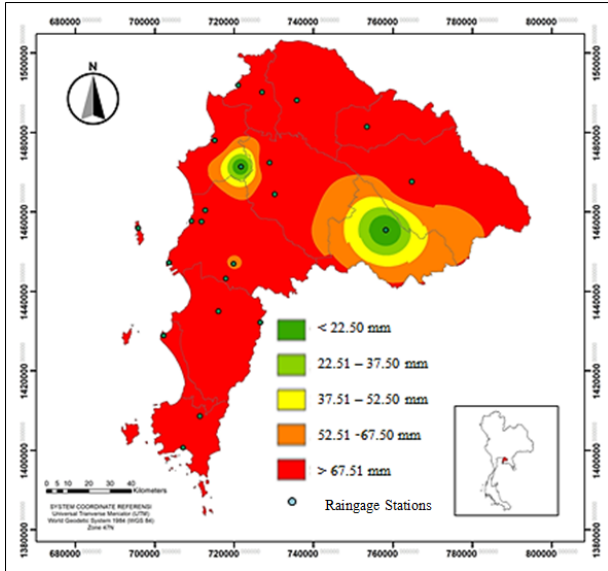


Fig. 2. Spatial average rainfall distribution by Inverse Distance Weighted Method, Choburi Province

Calibration and Verification on Weights of Input Factors

This study attempts to determine the best values of the weights of input factors W_1 to W_5 for the study area. This is done by calibration of the weights W 's of input data by comparing the computed matching index M which relates to the magnitude of flood inundation area with the satellite observed flood inundation area in the years 2013 and 2010. For the class 1 to class 5 of each input factor, the scale values R from 1 to 5 are assigned respectively. In the calibration, the weights of the input factors W_1 to W_5 are tried by changing from their assumed initial values that is $W_1=5$, $W_2=4$, $W_3=3$, $W_4=2$ and $W_5=1$. The number of combinations of W_1 to W_5 is equal to factorial 5 or $5 \times 4 \times 3 \times 2 \times 1 = 120$ combinations. Based on these combinations, 120 calibration runs of the ArcGIS were done for each year to obtain 120 values of matching index M from (1). To classify the matching index M , the full range of M from maximum and minimum possible values in ArcGIS is subdivided into five classes, e.g., 0-20% for very low, 21-40 % for low, 41-60% for moderate, 61-80% for high and 81-100% for very high. These classes were shown by different color codes for visualization in the plotted figures. In the same way, the observed flood inundation areas are classified into the same five classes. In each calibration run, the class of M and the class of satellite observed inundation area are compared for each grid. The score of agreement between the class

of M and the class of observed inundation is evaluated by using the following procedure.

The observed flooded area is divided into square grids each of 1 km by 1 km. By comparing the class of M and of the class of observed inundation in each grid, the score of agreement is counted as 1 if the agreement is perfect. If the agreement is not perfect, a score of agreement of less than 1 is read from the criteria in Table 2. The total score of agreement for the whole flooded area is the sum of all scores given to all grids. The overall average score of agreement of all grids in the flooded area is equal to the sum of scores of all grids divided by the total number of grids in the flooded area.

TABLE 2. SCORES FOR AGREEMENT BETWEEN CLASSE OF M AND CLASS OF SATELLITE OBSERVED DATA

Score of Agreement between M & Satellite		Classes of Calculated M				
		Very high	High	Mode rate	Low	Very low
Classes of Satellite Data	Very high	1	0.75	0.50	0.25	0
	High	0.75	1	0.75	0.50	0.25
	Mode rate	0.50	0.75	1	0.75	0.50
	Low	0.25	0.50	0.75	1	0.75
	Very low	0	0.25	0.50	0.75	1

IV. RESULTS OF MODEL CALIBRATION AND VERIFICATION

Model calibration is done for the floods in 2013 and 2010. The weights of the five input factors W_1 to W_5 obtained from their 120 combinations of W 's in each year are used. The agreement between the computed matching index M of the calibration run no. 9 and the satellite observed inundation data is found to yield the best comparison with $W_1=5$, $W_2=3$, $W_3=1$, $W_4=4$ and $W_5=2$. The score of agreement is 81% for 2013 and 72% for 2010 respectively. The comparison results for 2013 are shown in Figs. 3 and 4 respectively.

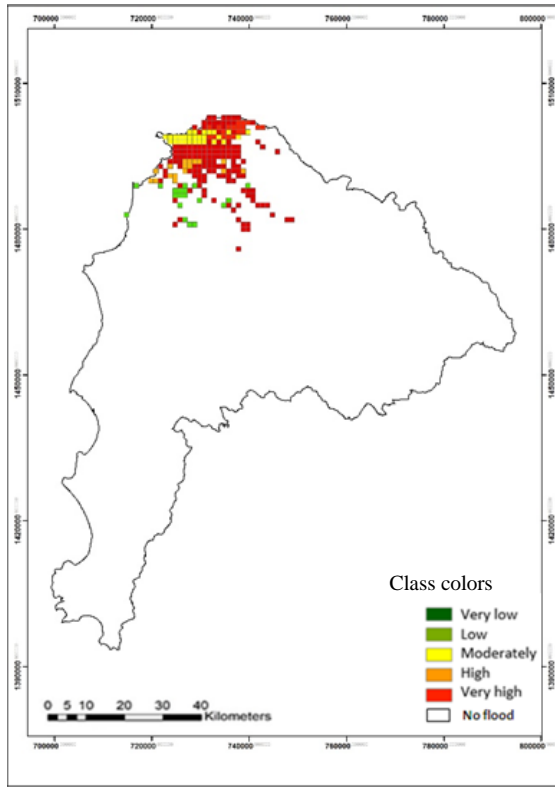


Fig. 3. Classes of satellite observed flood inundation area for model calibration in 2013

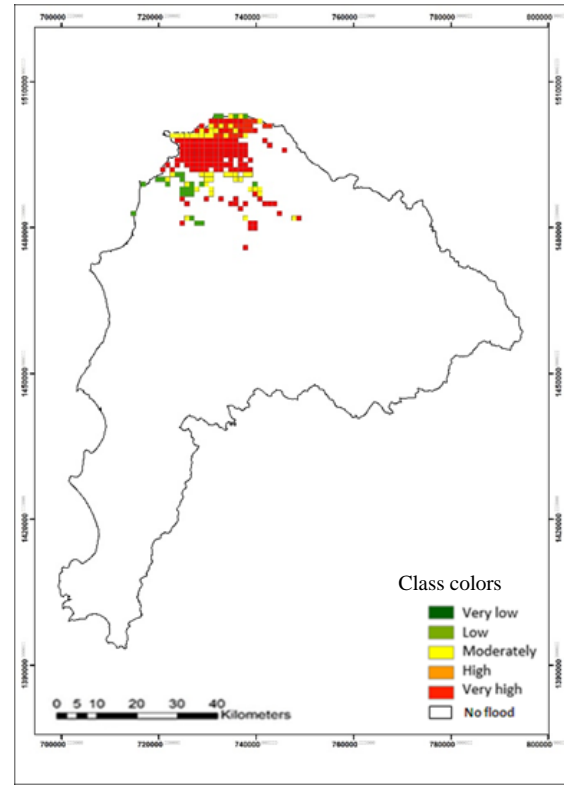


Fig. 4. Classes of computed matching index M for flood inundation area for model calibration in 2013

TABLE 3. WEIGHTS W's OF INPUT FACTORS IN CALIBRATION RUNS FOR 2013

Calib Run No.	% Agreement	Weights of Five Input factors				
		<i>W1 Rain</i>	<i>W2 Slope</i>	<i>W3 Elev</i>	<i>W4 Land use</i>	<i>W5 Soil type</i>
1	68 %	1	3	2	4	5
2	81 %	1	3	4	5	2
3	75 %	2	1	5	4	3
4	68 %	2	3	1	4	5
5	72 %	3	2	1	4	5
6	76 %	3	2	4	5	1
7	78 %	4	1	5	3	2
8	65 %	4	2	1	3	5
9 Best	81 %	5	3	1	4	2
10	44 %	5	1	4	2	3

For model verification, the best set of the weight W's obtained from calibration runs for 2013 and 2010 (case 9 in Table 3) is kept unchanged and is used to run the ArcGIS model for another big flood in 2008. Fig. 5 shows the classes of satellite observed inundation area in 2008 and Fig. 6 shows the classes of computed matching index M for flood inundation in the same year. Comparing both figures, the results show a satisfactory agreement with the percentage of agreement of 66%.

Hence, the successful model calibration and verification ensure the validity and reliability of the ArcGIS model developed in this study in applying it to predict annual flood inundation areas in other years.

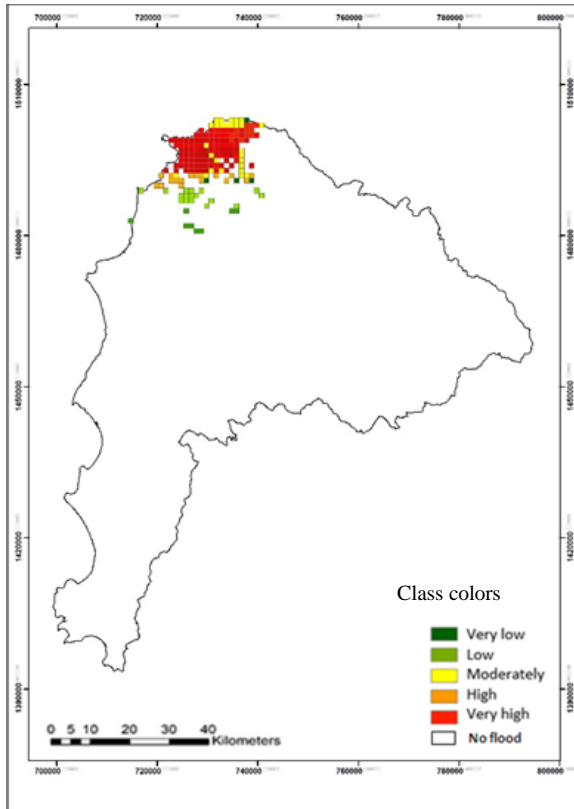


Fig. 5. Classes of satellite observed flood inundation area for model verification in 2008

V. DISCUSSIONS AND CONCLUSIONS

In other previous studies, the ArcGIS model was applied to determine flood hazard such as flood inundation areas by assuming the weights of various input factors that have influences on the floods. However, these weights were assumed rather intuitively without considering whether they correspond to the actual situation or not. In this study, an innovative approach is developed for the calibration and verification of the ArcGIS model for application in determining the flood inundation areas.

Usually, the flood inundation areas and flood depths depend on many input factors such as rainfall, ground slopes, ground elevation, land use types and soil types and other factors. The influence of these factors varies from one location to another location depending on the physiographical characteristics, hydro-meteorological characteristics and human activities in the study area.

The weights of these input factors in this study are assumed from 1 to 5, where 1 represents the weight for the lowest influence input factor and 5 for the highest influence. By using computer algorithm for successive iterations of weights of the inputs, the total number of iterations for the weights of the five inputs is equal

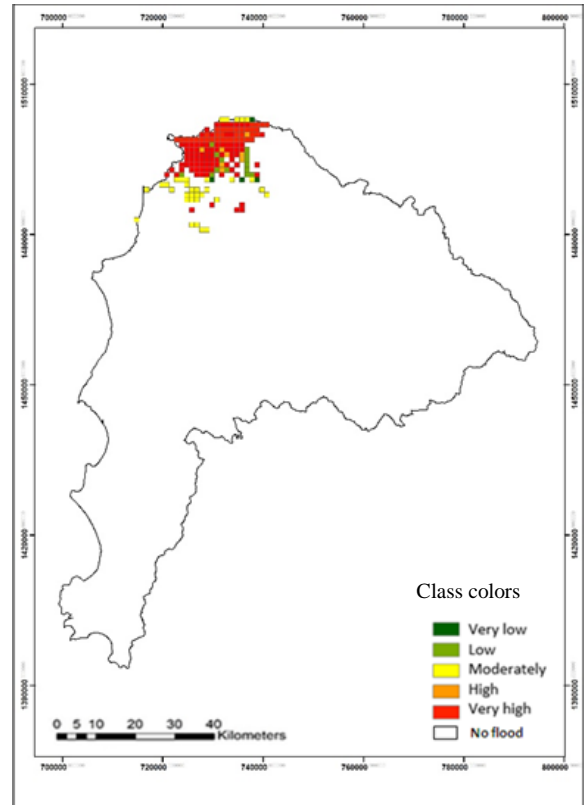


Fig. 6. Classes of computed matching index M for flood inundation area for model verification in 2008

to factorial 5 or 120 iterative computations. The 120 computations yielded 120 values of matching M for a selected flood. In this way, there are 120 calibration runs for each of the flood years of 2013 and 2010. Consistently a unique set of the integers $W_1=5$, $W_2=3$, $W_3=1$, $W_4=4$ and $W_5=2$ are found to be the best set for 2013 and 2010 (for calibration) and 2008 (for verification). From this result, it shows that rainfall is the most important input factor, followed by land use, ground slope, soil type and ground elevation respectively. The calibrated values of W_1 to W_5 obtained from the model calibration is found to work satisfactorily well and also in the verification runs of the year 2008. The assigned percentage ranges for classification of matching M and the observed inundation area in five categories described in Section III, work well for both calibration and verification in this study. However, these percentage ranges should be tested for other study areas depending on the physiographical and flood characteristics as well as the development of the study areas.

VI. RECOMMENDATIONS

The ArcGIS model requires calibration and verification of the weights of input data or input factors. With calibration and verification, the model output will be reliable. This is necessary to arrive at the best set of weights W 's which are applicable for estimation of flood hazard condition for different floods. With the

calibrated and verified set of the weights, the ArcGIS model and the developed methodology assures the validity and reliability in its application for estimation of flood hazard of other years in the study area. It is also recommended to apply this methodology for other areas.

ACKNOWLEDGMENT

The authors wish to express their sincere thanks to Sompop Kaewsawee of Kasetsart University, Sriracha Campus for providing useful data and information for this study.

REFERENCES

- [1] G.V. Bapulu, and R. Sinha, GIS in flood hazard mapping: a case study of Kosi River Basin, India. Noida: GIS Development. [ONLINE] Available from: http://home.iitk.ac.in/~rsinha/PDFs/2006_FloodGISdevelopment.pdf, 2005, [Accessed: 21/07/10].
- [2] Sompop Kaewsawee and Supatchaya Chuanpongpanich, "Flood hazard map and evaluation of economic loss in Chonburi Province", 3rd Technical Conference, Kasetsart University, Sriracha Campus, August, 2018, pp.235-243
- [3] Supapon Naka, Climate Change 2007: Analysis of flood risk area in the upper Pasak river basin, Lomsak, Phetchabun Province. Thesis in Natural Resources and Environment, Naresuan University, Thailand, 2015.
- [4] Nardnarade Akasuwan, Study on incidence of flood risk area factors using geographical information systems and prevention guidelines around western sub-basin area of Songkla lake in Phatthalung province. Master Thesis, Graduate School, Thaksin University, Thailand, 2009.
- [5] ESRI, *How weight overlay works*. [ONLINE] available at: <http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/how-weighted-overlay-works.htm>, 2016, [Accessed 10 August 2018].
- [6] D. A. Zimmer. 2011. "What is the Weighted Scoring Method." [ONLINE] Available at: <https://pmtterms.com/>, 2011, [Accessed 26 July 2018].
- [7] B. W. Ilbery. 1977. "Point Score Analysis: A Methodological Framework for Analysing the Decision-making Process in Agriculture". Journal of Economic and Social Geography, Vol. 68. Issue 2/ pp. 65-128.
- [8] ESRI, *How Inverse Distance Weight (IDW) Interpolation works*. April 26th 2018, Retrieved from ArcGIS 9.2 Desktop Help. [ONLINE] Available at: [http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=How_Inverse_Distance_Weighted_\(IDW\)_interpolation_works](http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=How_Inverse_Distance_Weighted_(IDW)_interpolation_works), 2007, [Accessed 18 July 2018].