Journal of Hydrology 595 (2021) 125655

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Explore training self-organizing map methods for clustering high-dimensional flood inundation maps

Li-Chiu Chang [a,*], Wu-Han Wang [a], Fi-John Chang [b]

[a] Department of Water Resources and Environmental Engineering, Tamkang University, New Taipei City 25137, Taiwan, ROC
[b] Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei 10617, Taiwan, ROC

ARTICLE INFO

This manuscript was handled by Andras Bardossy, Editor-in-Chief

Keywords:
Self-Organizing Map (SOM)
Flood inundation map
Topological map
Artificial Intelligence (AI)

ABSTRACT

The Self-Organizing Map (SOM) can supportively organize complex datasets such as highly dimensional flood inundation maps. Nevertheless, SOM may produce distinct patterns after being trained with identical samples or may not converge in clustering highly dimensional datasets, which causes usability concerns and prevents its applications from a broader spectrum. Motivated by such concerns, two training strategies (S1 and S2) were proposed to configure SOM based on a large number of highly dimensional flood inundation maps associated with two basins located in southern Taiwan. S1 focused mainly on the weights' adjustments in the ordering stage, while S2 would methodically balance the ordering and convergence activities on the weights' adjustments. The effectiveness and suitability of S1 and S2 were inspected in detail by using coverage ratio, flip detector, and five clustering indices based on their configured topological maps in the two basins. The clustering results showed that the flip detector and the coverage ratio could visibly and objectively examine the suitability of the configured topological map. It was noticed that the influences of the ordering and convergence stages upon both training strategies for building SOM could significantly affect the coverage ratio as well as flip condition. Comparing the SOM topological maps implemented separately with each strategy, S2 strategy has a lower probability of causing a flipping situation and takes far fewer iterations to train a model of the same network size, which indicates S2 is more efficient and effective than S1 in configuring the SOM topological map for representing regional flood inundation maps.

1. Introduction

In the face of increased flooding, disaster prevention is a crucial task for urbanized cities. Over the last decades, many flood simulation models have been developed to provide storm event-based urban flood inundation maps (Afshari et al., 2018; Cook and Merwade, 2009; Darabi et al., 2019; Dottori and Todini, 2011; Frank et al., 2012; Neal et al., 2012; Papaioannou et al., 2016; Rangari et al., 2018; Rong et al., 2020; Teng et al., 2017; Yamazaki et al., 2011; Yoo et al., 2013; Zhao et al., 2019). These models could grasp flood characteristics and provide simulated inundation maps of various designed or historical rainfall events, which are beneficial to city flood management. Nevertheless, despite monitoring data available on-line, they commonly failed to provide real-time regional inundation maps due to their heavy computation loads. Real-time regional flood forecasting is crucial for disaster prevention but remains a great challenge because of the complex interactions and disruptions engaging highly uncertain hydro-meteorological variables and the lack of high-resolution hydro-geomorphological data.

Artificial Intelligence (AI) has become one of the popular techniques in the studies of water resources management and flood forecasts in the last decades (ex., Chang et al., 2020; Nourani et al., 2014; Tikhamarine et al., 2020; Wang et al., 2009; Yaseen et al., 2015). Recent studies have indicated AI techniques could be a promising tool for flood simulation models to conquer the barrier of heavy computation loads and consistently provide real-time flood inundation maps based on monitored (or forecasted) rainfall during a storm event (ex., Berkhahn et al., 2019; Bui et al., 2016a, 2016b; Chang et al., 2014, 2018, 2019; Chapi et al., 2017; Falah et al., 2019; Rahmati et al., 2019; Shafizadeh-Moghadam et al., 2018). The self-organizing map (SOM) is a fascinating AI clustering tool for classifying highly dimensional datasets, where input vectors are projected onto a low-dimensional grid (map space) in an ordered fashion so that more similar vectors will be associated with nodes closer to the grid, whereas less similar vectors will be situated in nodes farther away

* Corresponding author.
E-mail address: changlc@mail.tku.edu.tw (L.-C. Chang).

https://doi.org/10.1016/j.jhydrol.2020.125655
Received 19 May 2020; Received in revised form 9 October 2020; Accepted 13 October 2020
Available online 18 October 2020

from the grid (Kohonen, 1990). SOM can project input vectors on prototypes of a two-dimensional regular grid to form a topological map that can be effectively utilized to explore the properties of the data. It has been broadly used in a wide variety of fields, ranging from industrial and financial domains (ex., Kessentini and Jeffers, 2018; Raptodimos and Lazakis, 2018) to water resources domains (ex., Adeloye et al., 2012; Chang et al., 2010, 2020; Cheng et al., 2018; Kalteh et al., 2008; Tsai et al., 2015, 2017), classification of satellite imagery data and rainfall estimation (Farzad and El-Shafie, 2017; Hong et al., 2006; Lin and Wu, 2007), rainfall-runoff modelling and analysis (Adeloye and Rustum, 2012; Chang and Tsai, 2016; Nourani and Parhizkar, 2013; Nourani et al., 2013; Srinivasulu and Jain, 2006), evapotranspiration (Adeloye

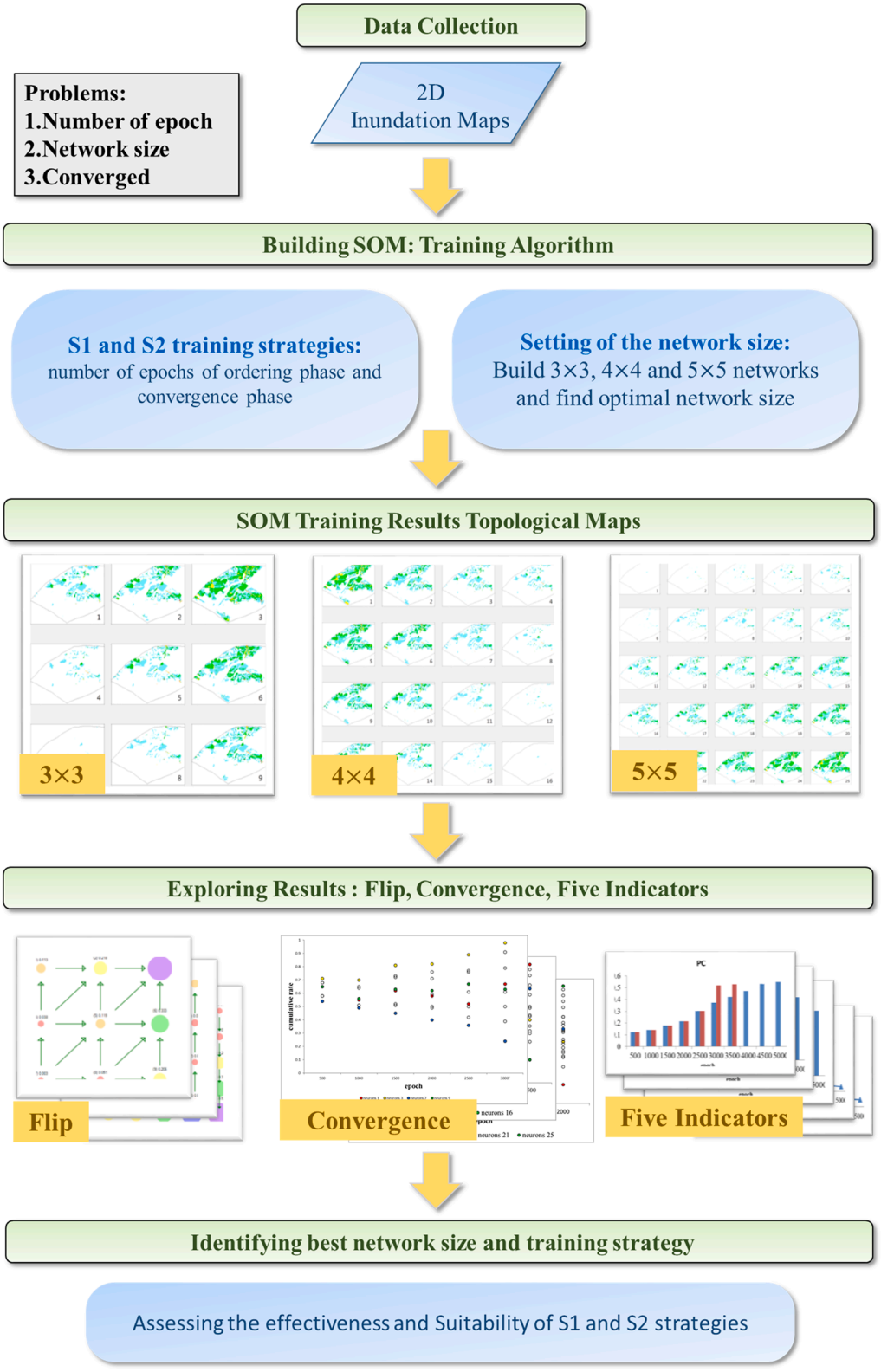


Fig. 1. The framework of this study.

et al., 2011), groundwater (Chang et al., 2016; Chen et al., 2018; Han et al., 2016; Haselbeck et al., 2019; Nakagawa et al., 2020; Nguyen et al., 2015), flood monitoring and management (Chang et al., 2007; Fotovatikhah et al., 2018; Kussul et al., 2011;), and regional flood inundation warning systems (Chang et al., 2014, 2018, 2019).

Notwithstanding SOM's wide applications, configuring a suitable SOM for highly dimensional datasets, such as a regional flood inundation map, is a formidable challenge. There are three common problems of configuring SOM, which are the twisting of the topological map, non-convergence, and the determination of the number of neurons (Aoki et al., 2009; Breard and Hamel, 2018). Besides, it could produce dissimilar patterns in clustering, even being trained by identical training datasets with the same parameters (Zhang et al., 2018). These issues do cause usability concerns and prevent its applications from a broader spectrum. Motivated by these practical concerns, we systematically explore the effectiveness and efficiency of two training strategies for configuring SOM as well as assess the suitability of the constructed SOM topological map using various evaluation indicators. We construct the topological map and analyze the clustering results of spatial–temporal distribution of regional flooding in two basins in Tainan City, Taiwan. We then present the results and evaluate the quality of the constructed topological maps encountered during model construction as well as provide further explanation of the proposed strategies assessed by evaluation indicators. Conclusions and suggestions are then given, which consolidates study results and main findings of this study and suggests possible research directions for future works.

2. Methods

This study proposes two training strategies to conquer three common problems (i.e., the twisting of the topological map, non-convergence, and the determination of the number of neurons) commonly encountered when configuring SOM based on a number of regional inundation maps for two river basins and explore the efficiency and suitability of the constructed SOM topological maps using various evaluation indicators. Fig. 1 shows the framework of the study, which contains data collection, SOM model construction, topological maps, exploration of results, assessment on the effectiveness of training strategies, and identification of the best network and training strategy. The methodology proposed to build SOM and evaluation indicators are described in the following sections.

2.1. Building SOM

SOM is commonly configured into a two-dimensional lattice that represents a grid-like structure among neurons (nodes) to form a topological map for tackling clustering problems and data exploration. Each neuron contains a vector with the same dimensionality of the input vector. The neurons are subsequently adjusted during the training phase based on competitive learning, where training is entirely data-driven, and the neurons of the map compete with each other (Alhoniemi et al., 1999). In every iteration, each training sample (input vector) is allocated to the neuron to which it best matches, denoted as the best match neuron (BMN), where the representative (central) vector of the matched neuron is updated so that the BMN moves closer to the input vector. Besides, the weights of BMN's neighboring neurons are methodically updated so that the surrounding neurons around the BMN are systematically adapted towards the training sample. After completing the training process, the final topology would show that the neurons on the grid become ordered, where similar neurons stay closer to each other. In contrast, less similar neurons remain farther away from each other. As a result, training samples are distributed across the constructed topological map in a similarity preserving way.

A fundamental question is whether the topological map can well organize and meaningfully interpret the data after training. To the best of our knowledge, the clustering results are subject to variability in initial conditions, convergence to local minima, and sampling complexity (Chang et al., 2014; Mzelikahle et al., 2017). We have experienced three significant problems when building the SOM topology, including the twisting map, the selection of the number of epochs, and the optimal network size. We propose a methodology to automatically configure SOM for establishing a meaningful topological map for a large number of highly dimensional flood inundation maps.

2.2. Mathematical formulation

A brief mathematical formulation of SOM, which is based mainly on Kohonen (1990), is shown as follows. The input vector X has N dimensions (Eq. (1)). Each neuron is represented by an N-dimensional weight vector (Eq. (2)). The best match neuron (BMN), winning neuron selection, is shown in Eq. (3). The topological neighborhood (i.e., the neighborhood of a neuron in a topological space) in the kth iteration is shown in Eq. (4). The weight adjustment amount Δw_j of the jth neuron in the adjacent area is shown in Eqs. (4)–(7), where Eq. (6) defines a time-varying neighborhood function around a winning neuron. The final weight is expressed as Eq. (8).

$$\mathrm{X} = [x_1, x_2, x_3 \cdots x_N]^{\mathrm{T}} \tag{1}$$

$$W_j = \left[w_{j1}, w_{j2} w_{j3} \ldots w_{jN}\right]^T, j = 1, 2, 3 \ldots, M \tag{2}$$

$$q(X) = \min_{\forall j} X - W_{j2}, j = 1, 2, 3 \ldots, M \tag{3}$$

$$\mathrm{R(k)} = R_0 \exp\left(-\frac{k}{\tau_1}\right) \mathrm{k} = 0, 1, 2, \cdots \tag{4}$$

$$\tau_1 = \frac{1000}{\log(R_0)} \tag{5}$$

$$h_{qj}(\mathrm{k}) = \exp\left(-\frac{\|r_j - r_q\|^2}{2R^2(k)}\right) \mathrm{k} = 0, 1, 2, \cdots \tag{6}$$

$$\Delta w_j = \eta(k) h_{q\,j}(k)\big(x(k) - w_j(k)\big) k = 0, 1, 2, \ldots \tag{7}$$

$$\mathrm{w_j}(k+1) = \mathrm{w_j}(k) + \Delta \mathrm{w_j k} = 0, 1, 2, \cdots \tag{8}$$

where M is the number of neurons, W_j is the weight vector of the jth neuron, R(k) represents the topological neighborhood in the k-th iteration,R_0 is the initial radius preset to cover the size of all neurons so that all neurons will fall within the range corrected by the neighborhood function,τ_1 is a constant value calculated by Eq. (5), $h_{qj}(\mathrm{k})$ denotes the amplitude of the topological neighborhood centered with the j-th neuron, r_j and r_q are the position of the j-th node and the q-th winner node in the topological map, respectively, and η is the learning rate.

2.3. Weight adjustment process

The initial weights of SOM are randomly generated, where the relationship between neurons is of extreme disorder; thus, the weight adjustment is the crucial process to configure a suitable topological map. The weight adjustment process is usually executed in two stages, i.e., ordering and convergence, where the ordering stage involves a rough training while the convergence stage is responsible for a finer training (Kohonen, 1990). In the ordering stage, the weight updating process will adjust the neurons of SOM so that the topological ordering of the weight vectors takes place. In the convergence stage, the weight updating process will adjust those neurons close to the best match neuron for fine-tuning the topological map. The number of iterations and the learning rates of these two stages would significantly affect the formation of the topology. Because the initial weights are in extreme disorder, the learning rate and neighborhood radius should be set as large values in

the first stage. To make the topological map gradually present the order, we have to tune the neurons so that neighboring neurons have similar weight vectors. The initial learning rate η_0 usually starts at 0.1, and then the learning rate, $\eta(k)$, decreases gradually as the number of iterations (k) increases, as shown in Eq. (9). $\tau 2$ is set as 1000.

$$\eta(k) = \eta_0 exp\left(-\frac{k}{\tau_2}\right) \quad k = 0, 1, 2, \ldots \tag{9}$$

In the second stage, smaller learning rate and neighborhood radius are favorable for fine-tuning the topological map. η_0 is changed to 0.01, and the range of neighborhood radius is also reduced so that only a few neurons staying close to the winner neuron would be fine-tuned.

We notice that the initial learning rate in the second (convergence) stage is much smaller than that of the first (ordering) stage, which is because the first stage is a rough training while the second stage is responsible for a finer training. After a number of updating (ex. 1000 epochs), the exponential-decay formulas shown in the neighborhood function (Eq. (6)) will reduce to a small value; thus, the weight adjustment (Eqs. (7) and (8)) would be effected only a few neighboring neurons around the winning neuron. The numbers of epochs for these two stages are crucial and must be set at first because such numbers could significantly influence the convergence as well as the suitability of the constructed topology.

2.4. Two training strategies

Two training strategies (S1 & S2) are proposed to investigate their effectiveness and suitability of SOM for constructing the flood inundation topologies of two river basins in Tainan City, Taiwan.

The first strategy, denoted as S1, intends to keep training the SOM network until no apparent changes occur in the weights of the neurons (i.e., coverage change in the 500th iteration is less than 5%) in the ordering stage. The training is transferred to the convergence stage, where the training stops if no obvious changes in the weight values can be made (i.e., coverage change in the 500th iteration is less than 5%).

The second strategy, denoted as S2, is to keep training the network until the coverage ratio of weights reaches 50% in the ordering stage. Then the weight adjustment process is transferred to the convergence stage and keeps on until no obvious changes in the weight values can be made (i.e., coverage change in the 500th iteration is less than 5%).

We notice that S1 is implemented mainly in the ordering stage so that the weights' adjustments in this stage will be conducted almost all the way to the end (less than 5%). In contrast, S2 would simultaneously balance the ordering and convergence activities in the weight updating process for roughly configuring the topological map and then fine-tuning the topological map.

2.5. Clustering evaluation indicators

A variety of quality measures developed over the years attempt to quantify how well the underlying data can be represented by a topological map of SOM. SOM can be trained with different map sizes to present the deviations of the data, and thus identifying the optimal size of SOM is crucial, and the first thing to decide. If the map size is too small, it will lose some important features that should be detected. Nevertheless, if the map size is too big, the differences between neurons (clusters) could be too little. A number of clustering evaluation indicators were proposed in the literature, which provided insights into the selection of a map size using quantitative indicators.

For instance, Srinivas et al. (2008) showed different clustering indicators for finding the optimal number of clusters in flood frequency analyses, and Farsadnia et al. (2014) combined SOM with three clustering methods to find the optimal number of clusters through different clustering indicators. Tian, et al. (2014) introduced the k-nearest neighbor algorithm to improve self-organizing maps for anomaly detection using healthy training data from experiments on cooling fan bearings.

Nevertheless, to the best of our knowledge, there is no general theoretical principle to determine the optimum map size, and there is no standard evaluation indicator created to evaluate the proximity between SOM neurons. Thus, the suitability of the formulated SOM topology should be explored by using different clustering evaluation indicators.

Here we compare quality measures by using five indicators to investigate the convergence and clustering effect (distinguishable the classification) of the SOM model. The principles and physical meanings of the five indicators are explained as follows.

I) The Partition Coefficient (PC) index, proposed by Bezdek (1981), shown in Eq. (10) presents the amount of overlap between clusters. Its membership degree is determined by the distance between the sample and the center of each cluster. The closer the sample to the cluster center is, the greater the membership degree is, and vice versa. Therefore, the larger the PC value is, the closer the samples are to cluster centers, which leads to a more obvious grouping effect. The membership degrees of each sample corresponding to different clusters is calculated by Eq. (10). μ in Eq. (11) can be set as any value, where the general setting is 2. The total membership degree should be 1.

$$\mathrm{PC} = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n}(u_{ij})^2 \tag{10}$$

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{\|x_j - c_i\|}{\|x_j - c_k\|}\right)^{\frac{2}{\mu-1}}} \quad \sum_{i=1}^{c} u_{ij} = 1 \tag{11}$$

where u_{ij} is the membership degree, c_i is the cluster center, x_j is the sample, n is the input dimension (the number of grids in the inundation map). c is the number of clusters.

II) The Classification Entropy (CE) is an index of the entropy principle, which calculated the fuzziness of the cluster partition (Farsadnia et al., 2014). When the probability of the event is more uniform, the entropy value is more unpredictable and larger. The CE value is calculated by Eq. (12). The smaller the entropy is, the smaller the CE is.

$$\mathrm{CE} = \frac{1}{-n\left[\sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}\log u_{ij}\right]} \tag{12}$$

III) The Xie–Beni (XB) indicator, proposed by Xie and Beni (1991), shown in Eq. (13) is a function of the data set and the centroids of the clusters, where the numerator is the deviation of the sample from the cluster center while the denominator is the smallest distance between different groups. A smaller value of the XB index implies a small intra-group deviation whereas a large distance between groups.

$$\mathrm{XB} = \frac{\sum_{i=1}^{c}\sum_{j=1}^{n}(u_{ij})^{\mu}\|x_j - c_i\|^2}{n \min\limits_{i \neq k}\|c_i - c_k\|^2} \tag{13}$$

IV) The Davies-Bouldin index (DBI), proposed by Davies and Bouldin (1979), shown in Eq. (14), is a metric built to express the performance of clustering tools and to determine the number of clusters. The DBI is similar to the XB indicator. The smaller the DBI is, the smaller the distance difference between the samples in

the group is. The larger the distance between the centers of different groups is, the more distinguishable the classification is.

$$DBI = \frac{1}{c}\sum_{i=1}^{c} \max_{i \neq j}\left(\frac{\overline{C_i} + \overline{C_j}}{\|c_i - c_j\|}\right) \quad (14)$$

where $\overline{C}$ is the average distance from all samples of this cluster to the center of this cluster.

V) The Silhouette Coefficient is created as a measure of cluster density and separation (Rousseeuw, 1987), which can evaluate a particular clustering of a dataset and compare it with other clusterings of the same dataset (Layton et al. 2013). The SC can be leveraged by measuring the correlation between the authorship distance method and the true authorship, evaluating the quality of the distance method. However, we show that the SC can be severely affected by outliers. The Silhouette Coefficient (SC) is shown in Eq. (15). S(i) ranges between −1 and 1. When the distance between groups is close, there is no much difference between the value of a(i) and b(i). Therefore, the value of S(i) tends to 0. When the average distance between the samples in the cluster and the cluster center (a) is greater than the center distance between the cluster center and the other clusters (b), S is a negative value, which means the classification is less suitable. Therefore, the larger the SC value is, the better the classification is.

$$S(i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}} \quad SC = \frac{1}{n}\sum_{i=1}^{n} S(i) \quad (15)$$

where $a(i)$ is the average distance between the i-th sample and the rest of the samples in the same cluster, and$b(i)$ is the minimum value from the set of average distance between the i-th sample to all samples in other clusters.

When the distance within the cluster is small, and the distance between the groups is large, the SC value is closer to 1. Thus, the larger the SC value is, the better the classification is.

3. Study areas and materials

Tainan City in southwestern Taiwan covers an area of about 2200 square kilometers, and its terrain comprises mainly alluvial plains. The city is located in the south of the Tropic of Cancer and belongs to the subtropical climate. Rainfall is usually induced by monsoons and typhoons, and more than 80% of the annual rainfall concentrates in the wet season (June–November). Typhoon-induced torrential rain or thunderstorms are prone to flooding. Luermen Creek and the Yenshui Creek in Tainan City are our study areas (Fig. 2), and their basic information is given in Table 1.

Because there is no real monitoring inundation map, we used simulation datasets (maps) as the data sources for training the SOM models. The simulation datasets of the "Taiwan Flooding Potential Map" were obtained from the Water Resources Planning and Research Institute, Water Resources Agency, Taiwan. The SOBEK (WL/Delft

Table 1
River information.

Basin	Area (km^2)	Length (km)	Slope
Luermen	42 (10128 grids)	5.93	1/2000–1/3000
Yenshui	343.74 (54751 grids)	41.3	1/295

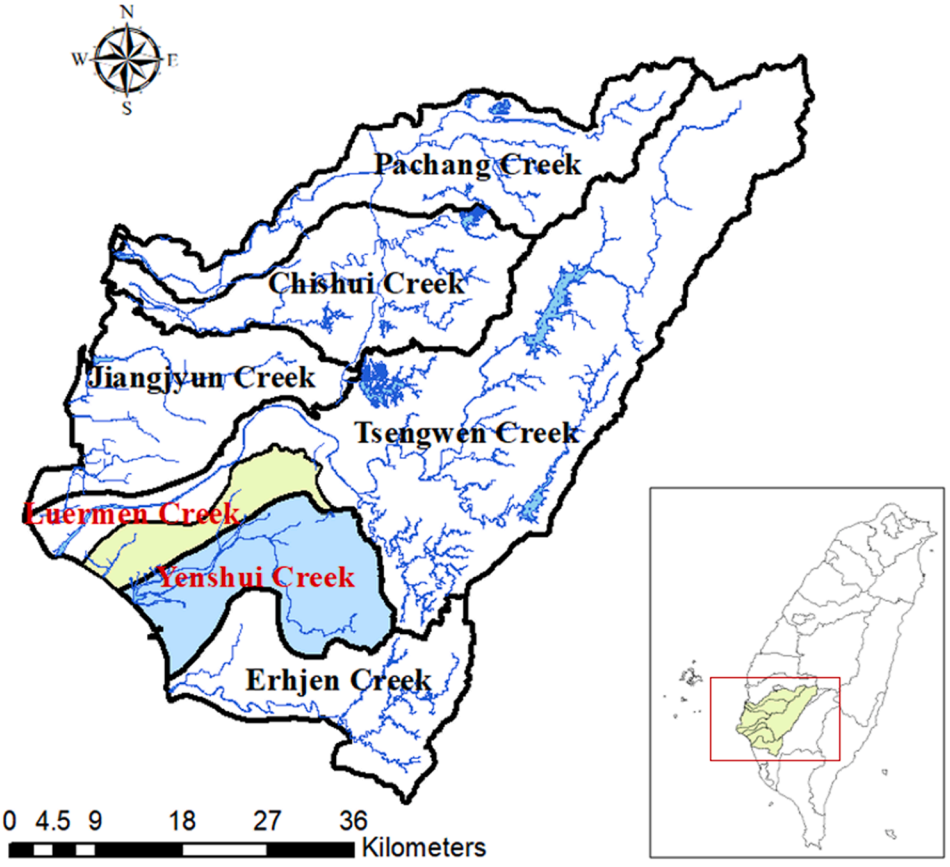


Fig. 2. Locations of Luermen Creek and Yenshui Creek in Tainan City of Taiwan. Source of the materials: "Tainan City Inundation Potential Map" from Water Resources Planning Institute, Water Resources Agency, Ministry of Economic Affairs.

Hydraulics, 1995), which had been used to simulate regional flood inundation maps in Taiwan (ex. Doong, et al., 2016; Yang et al., 2018), was used to simulate the flooding in the study areas to provide flood simulation datasets.

Twenty-seven scenarios were designed by a combination of three rainfall duration (6, 12, and 24 h) with nine rainfall amounts (100–800 mm) under various recurrences (i.e., 2- up to 500-year) of hydrological conditions. A total of 27 rainfall events (i.e. 3 durations × 9 rainfall amounts) constituted by various durations and rainfall amounts (Table 2) were used to simulate the corresponding flood inundation maps for both basins. The grid size of the inundation map is 40 × 40 m^2. We do not consider upland areas that have "never" been inundated (i.e., less than 10-mm inundation depth under any condition) so as to reduce the area as well as the input dimension for making computation less time-consuming. The Luermen Creek has 783 datasets (783 hourly inundation maps of 27 simulation scenarios), and each dataset has 10,128 input dimensions (grid's depths). The Yenshui Creek has 1539 datasets (1539 hourly inundation maps of 27 simulation scenarios), and each dataset has 54,751 input dimensions. For each basin, its datasets will be used as input vectors to configure a topological map as output. Thus, the SOM training algorithm is applied to cluster a large number of highly dimensional flood inundation maps (input vectors) and configure a visible topological map in both basins, individually.

4. Results and discussion

This study proposes two strategies (S1 & S2) to configure SOM for each study watershed and assesses their effectiveness and convergence of the constructed topological maps. Both strategies use the same set of initial weights for training; thereby, we explore their effects on topology convergence and investigate their flipping situations. The influences of the two strategies on building the SOM models are explored in detail through coverage ratio, flip detector, and five evaluation indicators of the clustering results. The coverage ratio represents the coverage of the maximum and minimum weights of neurons as a percentage of the training data. For example, if the maximum and minimum weights of SOM are 0.8 and 0.2, respectively, the coverage ratio is 60%. These measures allow us to assess the adequacy of the configured topological map quantitatively.

4.1. Explore the flip detector for the configured SOM

Fig. 3 presents a 4 × 4 topology (16 neurons) of the configured SOM, based on the 783 hourly inundation maps of the Luermen Creek basin, by using the S2 training strategy, where each circle represents one neuron, the number in parentheses represents the neuron number, and the number without parentheses represents the weight value of the neuron. We notice that the weight value of a neuron is the mean inundation depth of all the grids in the study area (ex. 10,128 grids in the Luermen). As shown, the weight value gradually increases from the lower right corner (neuron #16, 0.001 m) to the upper left corner (neuron #1, 0.503 m). A green arrow indicates the weight value increases in a correct direction, where the neighboring relationship is maintained correctly. While a red arrow represents the neighboring relationship is incorrect (flip), where the weight value increases in an incorrect direction. If that is the case, the topology needs to be retrained until the weight distribution is correct. In this study, a "flip" is identified if any two neighboring neuron values violate the rule of a correct direction, and their difference is significantly large, where a difference above 5% indicates a severe flipping while below 5% means an acceptable slight inversion. As shown in Fig. 3, the pair of neurons #2 and #6 violate the rule (red arrow), but the difference between their weights is only about 0.8%, indicating this is not a serious flipping situation and the difference in weight is within the acceptable range. Fig. 4 shows the configured 4 × 4 SOM topological map using the 783 hourly inundation maps. As shown, the topological map includes 16 neurons and each neuron represents an average inundation map of the inundation depths at 10,128 grids for the basin, where each grid denotes an area of 40 × 40 m^2. The average inundation map is obtained by computing the average of the number of regional inundation maps clustered in a neuron. As shown, neuron #1 has the largest regional inundation depths, where its average inundation depth is 0.503 m and a large portion of the grids have high inundation depth (even high than 2 m in some grids). In contrast, neuron #16 has the smallest regional inundation depths, where the average inundation depth is only 0.001 m and inundation does not occur in most of the grids).

The configured map has a suitable topological relationship where the regional inundation depths gradually increase from the lower right corner to the upper left corner.

4.2. Explore the efficiency of two training strategies

Both training strategies (S1 & S2) are used to construct the 3 × 3, 4 × 4, and 5 × 5 topological models for the Luermen Creek basin and the Yenshui Creek basin. The size and convergence of each constructed model are then assessed. The training epoch (iteration) for both training strategies is shown in Table 3. It appears it always takes fewer iterations to train a model of the same size by S2 than by S1, which clearly indicates S2 is more efficient than S1 in training the SOM models. We notice that the S1 strategy is based mainly on the ordering phase, which uses a large learning rate and neighborhood radius so that it is much easier to have a turn-over condition and result in a flipping phenomenon.

Table 4 presents the summarized results of the number of epochs and coverage ratio of SOM with three different sizes upon both training strategies (S1 & S2) in two phases (P1-ordering phase & P2-convergence phase) for both basins. The results indicate that (1) both strategies obtain similar final (P2) coverage ratios in the three SOM models for the two basins; (2) there is no much improvement (increasing the coverage) between P1 and P2 in respect of the S1 strategy, while there is the significant improvement between P1 and P2 in respect of the S2 strategy; (3) the number of epochs is much greater by using the S1 strategy than the S2 strategy; and (4) the total training epoch of the S1 strategy is much longer (larger) than that of the S2 strategy in all the cases. Thus, the S2 strategy is recommended, which trains the network in the ordering phase until the coverage ratio of weights reaches 50%, and then transfers the training to the convergence phase and continues training the weights until no obvious changes in the weight values are made.

Taking the 3 × 3 SOM model for the Luermen Creek basin as an example, Fig. 5 shows the cumulative ratio distribution at various numbers of iterations by both strategies (S1 & S2). The color-dots are the neurons of the 3 × 3 SOM. The nine vertical color lines are the average weight values of nine neurons corresponding to the cumulative rate curve (the blue curve), representing the cumulated average inundation

Table 2
Simulation scenarios of rainfall events.

Scenario Duration	1	2	3	4	5	6	7	8	9
6 h	100	150	200	250	300	350	400	450	500
12 h	150	200	250	300	350	400	450	500	550
24 h	200	250	300	350	400	450	500	650	800

Unit: mm.

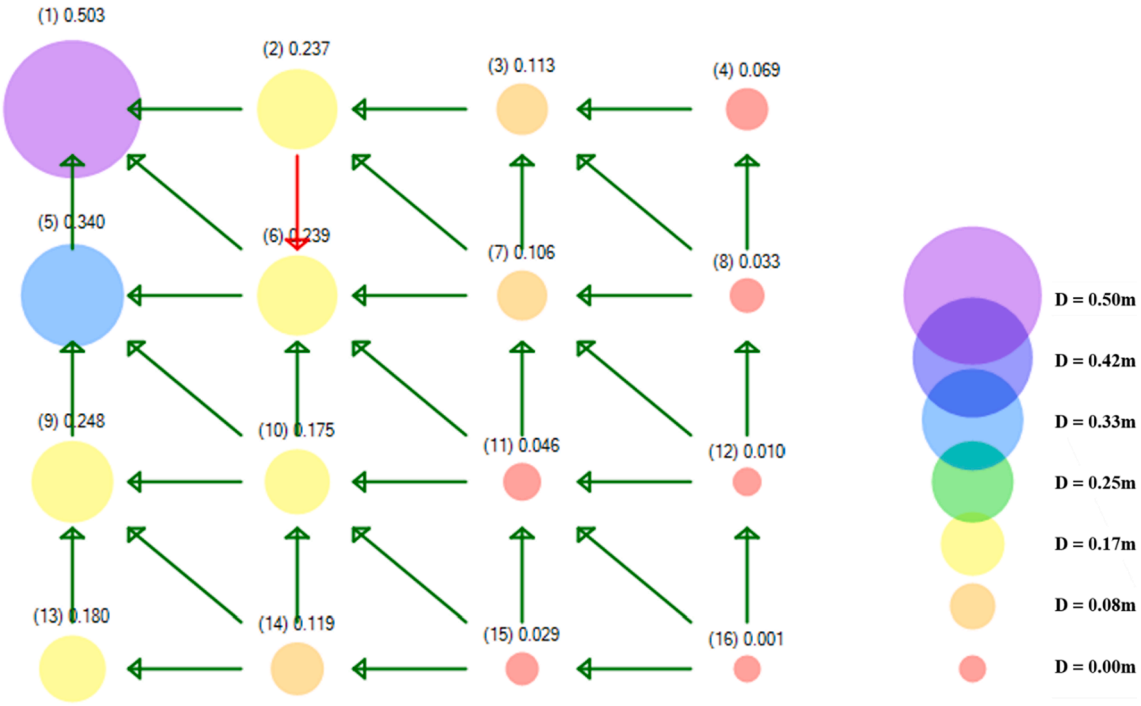


Fig. 3. Average inundation depth (unit: meter) in 16 neurons (4 × 4) of the Luermen Creek basin.

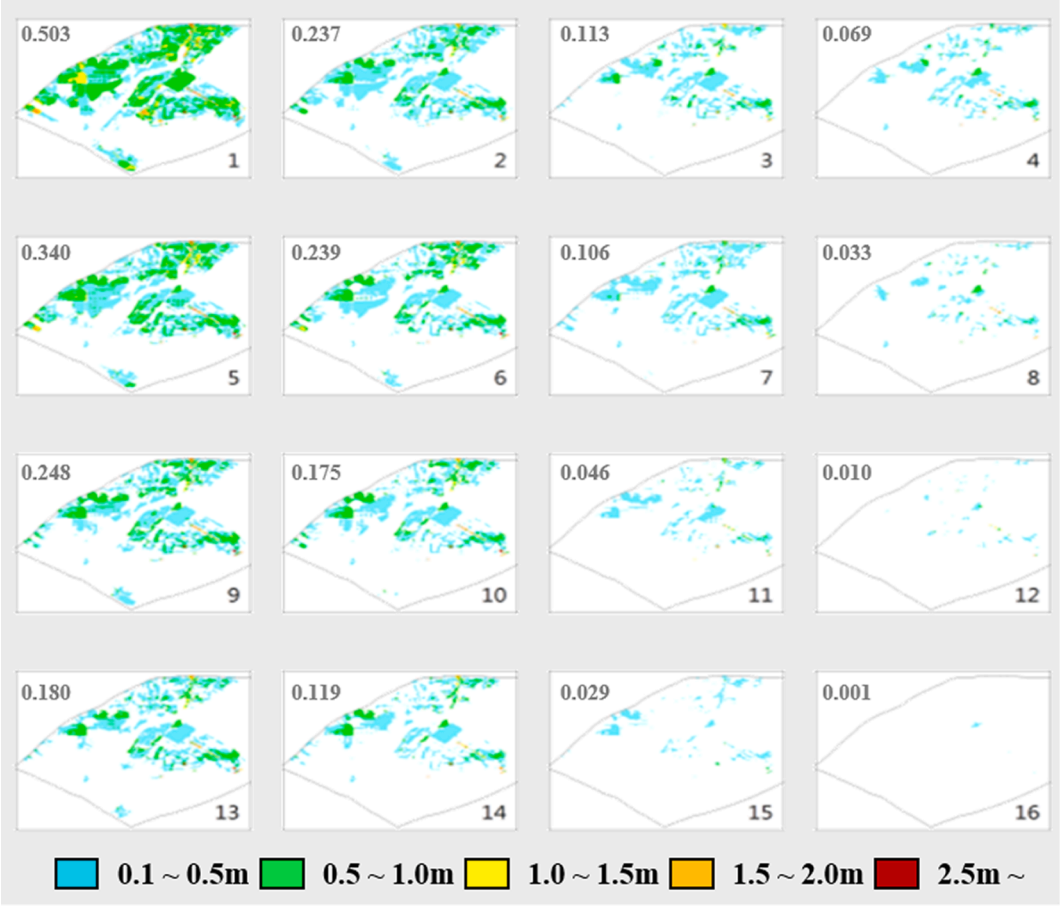


Fig. 4. The configured 4 × 4 SOM topological map of regional inudation depths in the Luermen Creek basin.

Table 3
Training time in respect to both training strategies (S1 and S2) for configuring the SOM.

Size of topological map	Training time (min)	3 × 3	4 × 4	5 x 5
Luermen Creek	S1	140	200	250
	S2	80	115	130
Yenshui Creek	S1	1400	2000	3100
	S2	930	1600	1800

depth of the counted datasets and the cumulative ratio of datasets (from zero inundation depth to the maximum inundation depth). It can be found that the neuron distribution of the two methods increases (expands) with the number of iterations (ex., 500 runs & 5000 runs of S1 and 500 runs & 3000 runs of S2). The initial weights concentrate in a small range (ex. from 0.55 to 0.72 in Fig. 5a) and then expand (ex. from 0.25 to 0.98 in Fig. 5b) gradually. The neurons with maximum and minimum cumulative rates in the training process fall on the two neurons farthest from the diagonal (neurons #3 and #7, respectively), where no other neurons surpass these two neurons, and their coverage ratios are 73% (0.25–0.98) for S1 and 74% (0.24–0.98) for S2. We notice that S1 requires 5000 runs (iterations) to train, while S2 only takes 3000 runs. The results indicate both strategies could provide almost the same coverage ratios, and S2 could reach the coverage ratio in a much shorter time than S1. Thus, S2 is much more efficient and effective than S1.

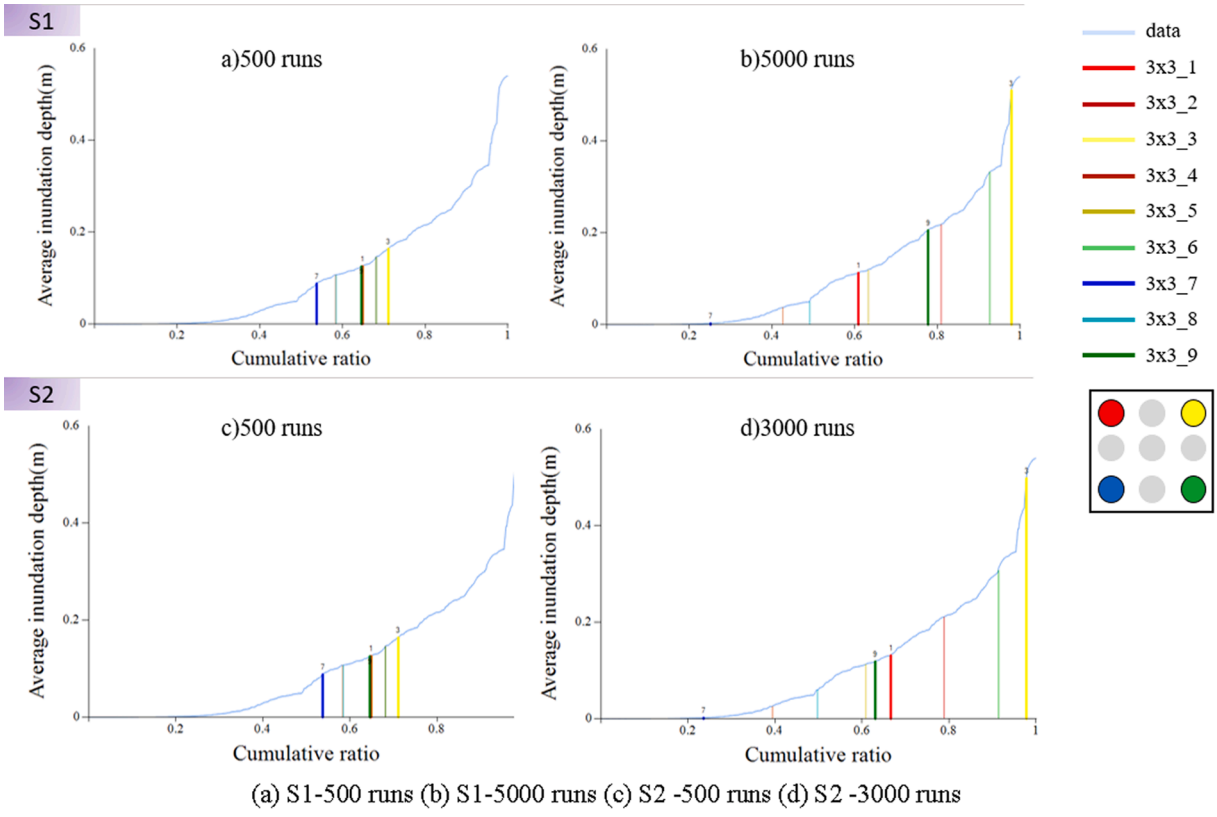
4.3. Examine the suitability of configured SOMs by evaluation indicators

Next, we explore the suitability of various sizes of SOM for clustering large numbers of inundation maps in the two study watersheds and assess their convergence. The five clustering evaluation indicators are implemented to evaluate the proximity (neighborhood relationship) between SOM neurons and investigate the convergence of the constructed SOM during the training process by both training strategies. As shown in Fig. 6, the values of PC, CE, and SC gradually increase (while the values of XB and DBI decrease) as the number of iterations increases. These results indicate that both training strategies can gradually

Table 4
Two training strategies (S1 and S2) implemented on the SOM models for the Luermen Creek basin and the Yenshui Creek basin.

Size of topological map			3 × 3		4 × 4		5 × 5	
Training stage			P1[a]	P2[b]	P1	P2	P2	P1
Luermen	S1	Number of epochs	4500	500	3500	500	3000	500
		Coverage ratio	73%	73%	78%	82%	79%	84%
	S2	Number of epochs	2500	500	2000	500	1500	500
		Coverage ratio	53%	74%	56%	79%	52%	83%
Yenshui	S1	Number of epochs	4000	500	2500	500	3000	500
		Coverage ratio	74%	77%	63%	73%	78%	80%
	S2	Number of epochs	2500	500	2000	500	1500	500
		Coverage ratio	55%	77%	53%	81%	52%	82%

[a] The ordering stage.
[b] The convergence stage.



(a) S1-500 runs (b) S1-5000 runs (c) S2 -500 runs (d) S2 -3000 runs

Fig. 5. The coverage ratio of nine neurons in the 3x3 SOM model along the iterations by both strategies (S1 & S2) for the Luermen Creek basin.

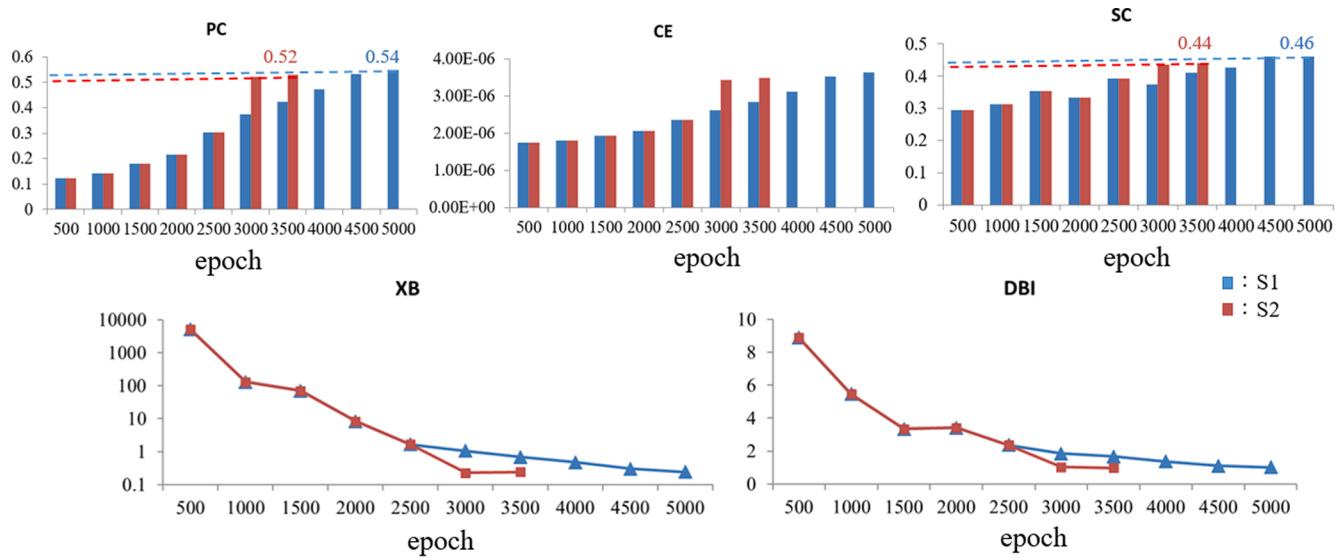


Fig. 6. Five clustering evaluation indicators by both strategies (S1 & S2) for the Luermen Creek basin.

distinguish (separate) the cluster centers during the training process, which also means the accuracy (coverage ratio) of classification gradually improves (increases). We notice that CE values are too small to be meaningful. Therefore, CE is considered unsuitable for identifying the classification of flood inundation maps. The value of SC does not change significantly along the training process, mainly because the SC indicator could not clearly identify whether the differences are within the group or between the groups due to SOM's neighboring relationship. Thus, the SC cannot be considered as an effective indicator for clustering the high dimensional flood inundation maps. We find the value of XB (as well as DBI) is relatively large at the initial iteration number (500 times) due to the randomly generated initial weights, where the weights between neurons are commonly very small. Therefore, the initial cluster centers (neuron weights) are very close to each other, and the classification between samples is not clear enough, resulting in considerably large values of XB and DBI. After thousands of iterations, the values of XB and DBI decrease significantly, which means these indicators are valid.

Here, we further explore the suitability of various sizes of SOM based on coverage ratios. Table 4 presents the coverage ratios corresponding to three SOM sizes, where a gradual increase in coverage ratio can be found as the map size increases. We notice that the coverage ratio corresponding to the 3 × 3 map size is relatively small (less than 77%), and thus the 3 × 3 map size is not considered applicable, taking the Yenshui Creek basin as an example. Fig. 7 shows the cumulative ratio for both of Luermen Creek basin and the Yenshui Creek. It indicates that the coverage ratio corresponding to 5 × 5 map is slightly higher than that of the 4 × 4 map (82% vs. 81% for Yenshui; and 83% vs. 79% for Luermen). Nevertheless, the cumulative ratios of the neurons in 5 × 5 models stay

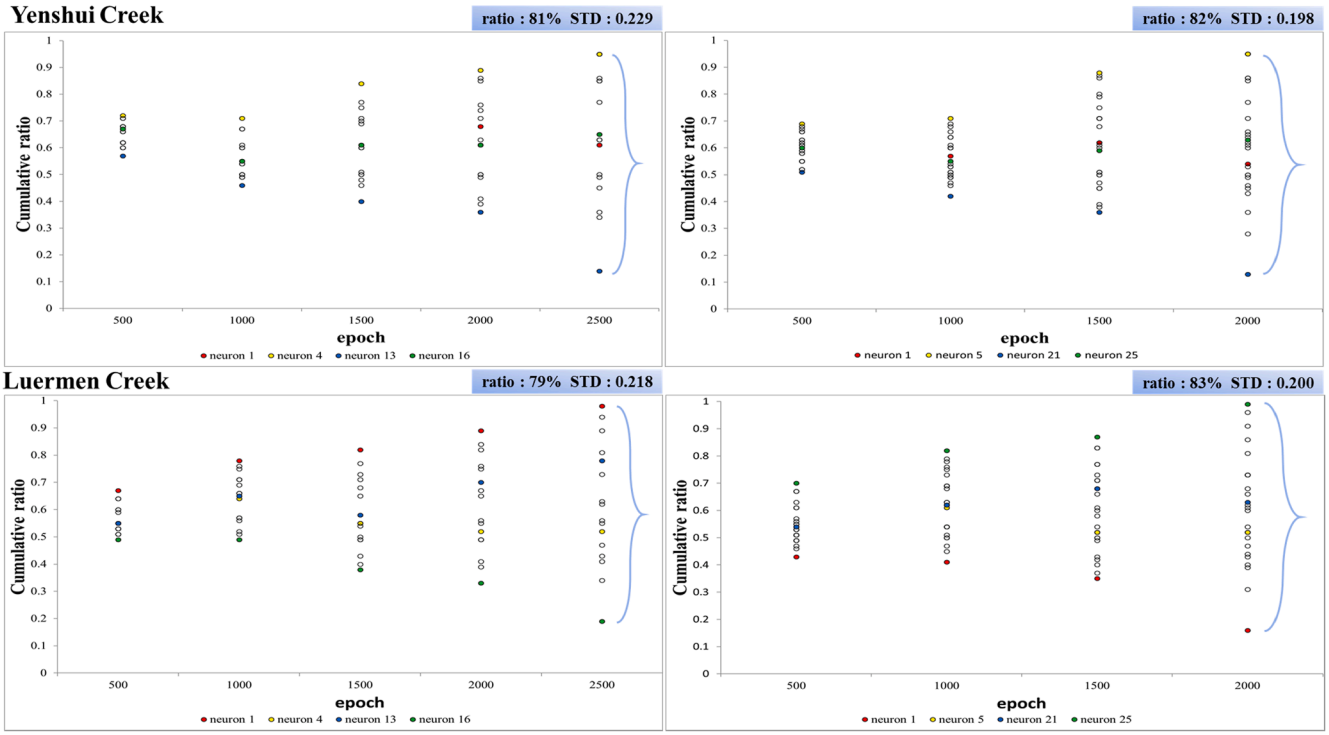


Fig. 7. The standard deviation (STD) and the coverage ratio of 4x4 and 5x5 SOM map sizes along the iterations by S2.

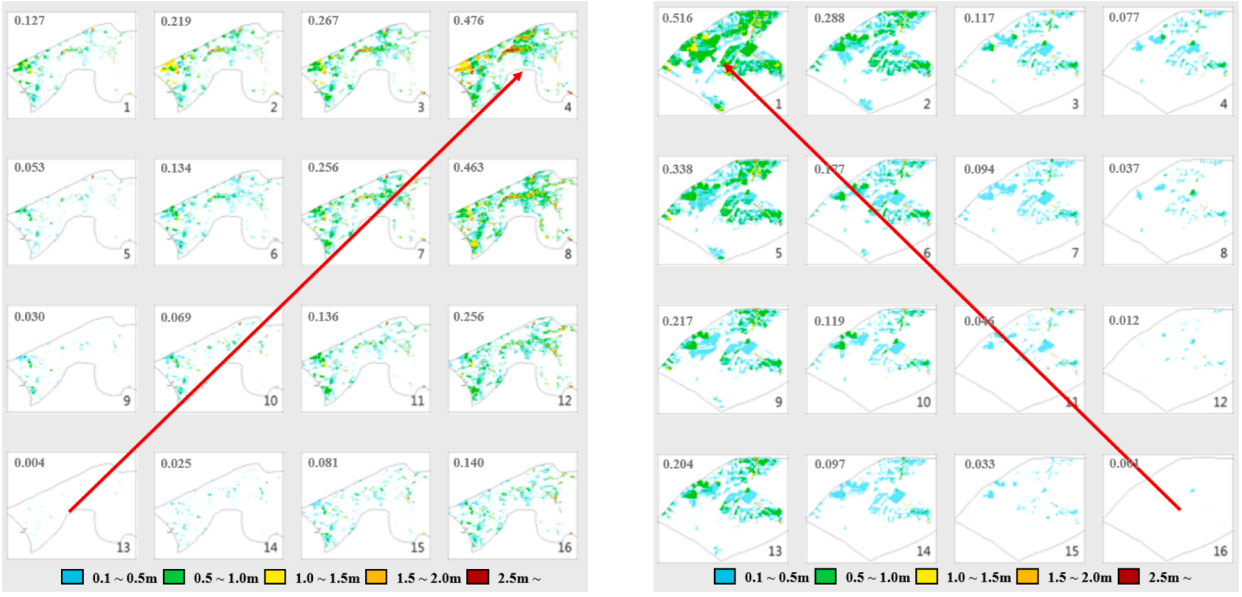
close together, and its standard deviation (STD) at 2000 iterations is small (0.198 for Yenshui; and 0.200 for Luermen), which suggests the model might encounter over-fitting. In contrast, the 4 × 4 model has a larger STD (0.229 for Yenshui; and 0.218 for Luermen) and more dispersed cumulative ratios at 2500 iterations. Thus we consider the 4 × 4 topography is the best choice for both basins. Fig. 8 presents the constructed 4 × 4 topological maps of both basins, where a continuously gradual weight change along the diagonal neurons and the nice proximity among neighboring neurons could be visibly identified. SOM with 4 × 4 neurons provided the best mixture of classification presentation and track neuron numbers in clusters producing a typology supporting our analysis. The benefits of this topological map are the ability to meaningfully categorize a large number of observed or simulated inundation maps as well as the visibility of the topological map to easily and visibly identify the best match inundation map according to the estimated and/or monitored hydrological condition.

5. Conclusions

SOM has been used extensively as an analytical and visualization tool in exploratory data analysis. The constructed topology could be visualized to give an insight into the topographic relationships of high-dimensional flood inundation maps. Thus, SOM could be a very promising application to regional flood warning systems if it is integrated with other machine learning models based on regional rainfall patterns such as the way how SOM was applied in the previous studies (Chang et al., 2014, 2018). Nevertheless, SOM may produce different patterns and/or non-convergence after training, and thus causes usability concerns. In this study, we propose two training strategies (S1 and S2) to train SOM models using a large number of highly dimensional flood inundation maps of two basins located in southern Taiwan and investigate their effectiveness and suitability of the constructed SOM models. We notice that S1 focuses mainly on the weights' adjustments in the ordering stage to almost all the way to the end, while S2 would simultaneously balance the ordering and convergence activities on the weights' adjustments. We examine the coverage ratio and the flipping of the topological map when the SOM model is applied to a large number of high-dimensional regional flood inundation maps. The main contributions of the proposed training strategies and evaluation indicators of the constructed SOM are four-fold and summarized as follows.

(1) Comparing the two training strategies for constructing SOM models for the two study basins, S2 has a lower probability of causing a flipping situation and takes much less training time (fewer iterations) to build a model of the same size than by S1, which clearly indicates S2 is more efficient than S1 in training the SOM models.
(2) The analytical results of the three constructed topological maps (i.e., 3 × 3, 4 × 4, and 5 × 5) for both basins indicate that the coverage ratio of the 3 × 3 map size is relatively low, whereas the 5 × 5 map size has some neurons with very similar accumulation ratios and result in an over-fitting situation. The 4 × 4 map size has a larger variation and well dispersed cumulative coverage ratios. Thus we consider the 4 × 4 topography is the best choice for both basins.
(3) The flip detector implemented in this study could visibly present and exam the suitability of the configured topological map. The computed coverage ratio along the training process could clearly indicate the suitability of the SOM map sizes and objectively assess the suitability of the topological map for a study area.
(4) Among the five indicators, PC, XB, and DBI can clearly express the training process as effective classification. The CE values are too small to compare, and the SC cannot well display the SOM map for the highly dimensional regional flood inundation applications. Thus CE and SC would not be recommended as effective evaluation indicators for assessing SOM's classification of regional flood maps. In short, PC, XB, and DBI can show the training of SOM is effective in categorization. However, CE and SC could not show effectiveness.

SOM is capable of clustering and data mining and is widely used in various hydrological processes. Configuring SOM could face issues of twisting topological maps and non-convergence, which prevents SOM applications from a broader spectrum. This study explores two weight adjustment strategies (S1 & S2) in detail, and the results suggest S2 strategy, which methodically balances the ordering and convergence activities on weight adjustment, could more effectively and efficiently construct a suitable topological map than the S1 strategy. Thus, S2 training strategy could be applied to effectively clustering a large number of highly dimensional datasets into a meaningful topological map for visibly explaining datasets in a wider spectrum.



(a) Yenshui Creek (b) Luermen Creek

Fig. 8. The constructed 4x4 topological maps of (a) Yenshui Creek basin and (b) Luermen Creek basin.

CRediT authorship contribution statement

Li-Chiu Chang: Conceptualization, Methodology, Supervision, Writing - original draft. **Wu-Han Wang:** Software, Methodology, Formal analysis, Investigation. **Fi-John Chang:** Conceptualization, Writing - original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was partly supported by the Water Resources Agency (WRA), Mistry of Economic Affairs, Taiwan (MOEAWRA1090316), and the Mistry of Science and Technology, Taiwan (Most107-2313-B-032-001). We thank the WRA for providing access to their data.

References

Adeloye, A.J., Rustum, R., 2012. Self-organising map rainfall-runoff multivariate modelling for runoff reconstruction in inadequately gauged basins. Hydrol. Res. 43 (5), 603–617.

Adeloye, A.J., Rustum, R., Kariyama, I.D., 2011. Kohonen self-organizing map estimator for the reference crop evapotranspiration. Water Resour. Res. 47 (8), W08523. https://doi.org/10.1029/2011WR010690.

Adeloye, A.J., Rustum, R., Kariyama, I.D., 2012. Neural computing modeling of the reference crop evapotranspiration. Environ. Modell. Software 29 (1), 61–73.

Afshari, S., Tavakoly, A.A., Rajib, M.A., Zheng, X., Follum, M.L., Omranian, E., Fekete, B. M., 2018. Comparison of new generation low-complexity flood inundation mapping tools with a hydrodynamic model. J. Hydrol. 556, 539–556.

Alhoniemi, E., Hollmén, J., Simula, O., Vesanto, J., 1999. Process monitoring and modeling using the self-organizing map. Integr. Comput. Aided Eng. 6 (1), 3–14.

Aoki, T., Ota, K., Kurata, K., Aoyagi, T., 2009. Ordering process of self-organizing maps improved by asymmetric neighborhood function. Cogn. Neurodyn. 3 (1), 9–15.

Berkhahn, S., Fuchs, L., Neuweiler, I., 2019. An ensemble neural network model for real-time prediction of urban floods. J. Hydrol. 575, 743–754.

Bezdek, J.C., 1981. Models for pattern recognition. In: Pattern Recognition with Fuzzy Objective Function Algorithms. Springer, Boston, MA, pp. 1–13.

Breard, G., Hamel, L., 2018. Evaluating self-organizing map quality measures as convergence criteria. Proceedings of the 2018 International Conference on Data Science, Las Vegas, Nevada, 86–92.

Bui, D.T., Pradhan, B., Nampak, H., Bui, Q.T., Tran, Q.A., Nguyen, Q.P., 2016. Hybrid artificial intelligence approach based on neural fuzzy inference model and metaheuristic optimization for flood susceptibility modeling in a high-frequency tropical cyclone area using GIS. J. Hydrol. 540, 317–330.

Chang, F.J., Chang, L.C., Wang, Y.S., 2007. Enforced self-organizing map neural networks for river flood forecasting. Hydrol. Process. 21, 741–749.

Chang, F.J., Chang, L.C., Kao, H.S., Wu, G.R., 2010. Assessing the effort of meteorological variables for evaporation estimation by self-organizing map neural network. J. Hydrol. 384, 118–129.

Chang, F.J., Tsai, M.J., 2016. A nonlinear spatio-temporal lumping of radar rainfall for modeling multi-step-ahead inflow forecasts by data-driven techniques. J. Hydrol. 535, 256–269.

Chang, F.J., Chang, L.C., Huang, C.W., Kao, I.F., 2016. Prediction of monthly regional groundwater levels through hybrid soft-computing techniques. J. Hydrol. 541, 965–976.

Chang, L.C., bin Mat Amin, I.M.Z., Yang, S.N., Chang, F.J., 2018. Building ANN-based regional multi-step-ahead flood inundation forecast models. Water 10 (9), 1283.

Chang, L.C., Chang, F.J., Yang, S.N., Kao, I., Ku, Y.Y., Kuo, C.L, bin Mat Admin, I.M.Z., 2019. Building an intelligent hydroinformatics integration platform for regional flood inundation warning systems.

Chang, L.C., Chang, F.J., Yang, S.N., Tsai, F.H., Chang, T.H., Herricks, E.E., 2020. Self-organizing maps of typhoon tracks allow for flood forecasts up to two days in advance. Nat. Commun. 11, 1983. https://doi.org/10.1038/s41467-020-15734-7.

Chang, L.C., Shen, H.Y., Chang, F.J., 2014. Regional flood inundation nowcast using hybrid SOM and dynamic neural networks. J. Hydrol. 519, 476–489.

Chapi, K., Singh, V.P., Shirzadi, A., Shahabi, H., Bui, D.T., Pham, B.T., Khosravi, K., 2017. A novel hybrid artificial intelligence approach for flood susceptibility assessment. Environ. Modell. Software 95, 229–245.

Chen, I.T., Chang, L.C., Chang, F.J., 2018. Exploring the spatio-temporal interrelation between groundwater and surface water by using the self-organizing maps. J. Hydrol. 556, 131–142.

Cheng, S.T., Tsai, W.P., Yu, T.C., Herricks, E.E., Chang, F.J., 2018. Signals of stream fish homogenization revealed by AI-based clusters. Sci. Rep. 8 (1), 15960.

Cook, A., Merwade, V., 2009. Effect of topographic data, geometric configuration and modeling approach on flood inundation mapping. J. Hydrol. 377 (1–2), 131–142.

Darabi, H., Choubin, B., Rahmati, O., Haghighi, A.T., Pradhan, B., Kløve, B., 2019. Urban flood risk mapping using the GARP and QUEST models: a comparative study of machine learning techniques. J. Hydrol. 569, 142–154.

Davies, D.L., Bouldin, D.W., 1979. A cluster separation measure. IEEE Trans. Pattern Anal. Mach. Intell. 2, 224–227.

Doong, D.J., Lo, W., Vojinovic, Z., Lee, W.L., Lee, S.P., 2016. Development of a new generation of flood inundation maps—a case study of the coastal city of Tainan, Taiwan. Water 8 (11), 521.

Dottori, F., Todini, E., 2011. Developments of a flood inundation model based on the cellular automata approach: testing different methods to improve model performance. Phys. Chem. Earth 36 (7–8), 266–280.

Falah, F., Rahmati, O., Rostami, M., Ahmadisharaf, E., Daliakopoulos, I.N., Pourghasemi, H.R., 2019. Artificial neural networks for flood susceptibility mapping in data-scarce urban areas. In: Spatial Modeling in GIS and R for Earth and Environmental Sciences. Elsevier, pp. 323–336.

Farsadnia, F., Kamrood, M.R., Nia, A.M., Modarres, R., Bray, M.T., Han, D., Sadatinejad, J., 2014. Identification of homogeneous regions for regionalization of watersheds by two-level self-organizing feature maps. J. Hydrol. 509, 387–397.

Farzad, F., El-Shafie, A.H., 2017. Performance enhancement of rainfall pattern–water level prediction model utilizing self-organizing-map clustering method. Water Resour. Manage. 31 (3), 945–959.

Fotovatikhah, F., Herrera, M., Shamshirband, S., Chau, K.W., Faizollahzadeh Ardabili, S., Piran, M.J., 2018. Survey of computational intelligence as basis to big flood management: challenges, research directions and future work. Eng. Appl. Comput. Fluid Mech. 12 (1), 411–437.

Frank, E., Sofia, G., Fattorelli, S., 2012. Effects of topographic data resolution and spatial model resolution on hydraulic and hydro-morphological models for flood risk assessment. Flood Risk Assessment and Management. Flood Risk Assess. Manage. https://doi.org/10.2495/978-1-84564-646-2/03.

Han, J.C., Huang, Y., Li, Z., Zhao, C., Cheng, G., Huang, P., 2016. Groundwater level prediction using a SOM-aided stepwise cluster inference model. J. Environ. Manage. 182, 308–321.

Haselbeck, V., Kordilla, J., Krause, F., Sauter, M., 2019. Self-organizing maps for the identification of groundwater salinity sources based on hydrochemical data. J. Hydrol. 576, 610–619.

Hong, Y., Chiang, Y.M., Liu, Y., Hsu, K.L., Sorooshian, S., 2006. Satellite-based precipitation estimation using watershed segmentation and growing hierarchical self-organizing map. Int. J. Remote Sens. 27 (23), 5165–5184.

Kalteh, A.M., Hjorth, P., Berndtsson, R., 2008. Review of the self-organizing map (SOM) approach in water resources: analysis, modelling and application. Environ. Modell. Software 23 (7), 835–845.

Kessentini, M., Jeffers, E., 2018. Visual exploration and analysis of bank performance using self organizing map. International Conference on the Sciences of Electronics, Technologies of Information and Telecommunications, December, 2018, Springer, Cham, pp. 420–434.

Kohonen, T., 1990. The self-organizing map. Proc. IEEE 78 (9), 1464–1480.

Kussul, N., Shelestov, A., Skakun, S., 2011. Flood monitoring from SAR data. In: Use of Satellite and In-Situ Data to Improve Sustainability. Springer, Dordrecht, pp. 19–29.

Layton, R., Watters, P., Dazeley, R., 2013. Automated unsupervised authorship analysis using evidence accumulation clustering. Nat. Lang. Eng. 19 (1), 95.

Lin, G.F., Wu, M.C., 2007. A SOM-based approach to estimating design hyetographs of ungauged sites. J. Hydrol. 339 (3–4), 216–226.

Mzelikahle, K., Mapuma, D.J., Hlatywayo, D.J., Trimble, J., 2017. Optimisation of self organising maps using the bat algorithm. Am. J. Inf. Sci. Comput. Eng. 3 (6), 77–83.

Nakagawa, K., Yu, Z.Q., Berndtsson, R., Hosono, T., 2020. Temporal characteristics of groundwater chemistry affected by the 2016 Kumamoto earthquake using self-organizing maps. J. Hydrol. 582, 124519.

Neal, J., Villanueva, I., Willis, T., Fewtrell, T., Bates, P., 2012. How much physical complexity is needed to model flood inundation? Hydrol. Process. 26 (15), 2264–2282.

Nguyen, T.T., Kawamura, A., Tong, T.N., Nakagawa, N., Amaguchi, H., Gilbuena, R., 2015. Clustering spatio–seasonal hydrogeochemical data using self-organizing maps for groundwater quality assessment in the Red River Delta, Vietnam. J. Hydrol. 522, 661–673. https://doi.org/10.1016/j.jhydrol.2015.01.023.

Nourani, V., Baghanam, A.H., Adamowski, J., Gebremichael, M., 2013. Using self-organizing maps and wavelet transforms for space–time pre-processing of satellite precipitation and runoff data in neural network based rainfall–runoff modeling. J. Hydrol. 476, 228–243.

Nourani, V., Baghanam, A.H., Adamowski, J., Kisi, O., 2014. Applications of hybrid wavelet–artificial intelligence models in hydrology: a review. J. Hydrol. 514, 358–377.

Nourani, V., Parhizkar, M., 2013. Conjunction of SOM-based feature extraction method and hybrid wavelet-ANN approach for rainfall–runoff modeling. J. Hydroinf. 15 (3), 829–848.

Papaioannou, G., Loukas, A., Vasiliades, L., Aronica, G.T., 2016. Flood inundation mapping sensitivity to riverine spatial resolution and modelling approach. Nat. Hazards 83 (1), 117–132.

Rahmati, O., Darabi, H., Haghighi, A.T., Stefanidis, S., Kornejady, A., Nalivan, O.A., Bui, D.T., 2019. Urban flood hazard modeling using self-organizing map neural network. Water 11 (11), 2370.

Rangari, V.A., Gonugunta, R., Umamahesh, N.V., Patel, A.K., Bhatt, C.M., 2018. 1D-2D modeling of urban floods and risk map generation for the part of Hyderabad city. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.

Raptodimos, Y., Lazakis, I., 2018. Using artificial neural network-self-organising map for data clustering of marine engine condition monitoring applications. Ships Offshore Struct. 13 (6), 649–656.
Rong, Y., Zhang, T., Zheng, Y., Hu, C., Peng, L., Feng, P., 2020. Three-dimensional urban flood inundation simulation based on digital aerial photogrammetry. J. Hydrol. 584, 124308.
Rousseeuw, P., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20, 53–65.
Shafizadeh-Moghadam, H., Valavi, R., Shahabi, H., Chapi, K., Shirzadi, A., 2018. Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping. J. Environ. Manage. 217, 1–11.
Srinivasulu, S., Jain, A., 2006. A comparative analysis of training methods for artificial neural network rainfall–runoff models. Appl. Soft Comput. 6 (3), 295–306.
Srinivas, V.V., Tripathi, S., Rao, A.R., Govindaraju, R.S., 2008. Regional flood frequency analysis by combining self-organizing feature map and fuzzy clustering. J. Hydrol. 348, 148–166.
Teng, J., Jakeman, A.J., Vaze, J., Croke, B.F., Dutta, D., Kim, S., 2017. Flood inundation modelling: a review of methods, recent advances and uncertainty analysis. Environ. Modell. Software 90, 201–216.
Tian, J., Azarian, M.H., Pecht, M., 2014. Anomaly detection using self-organizing maps-based k-nearest neighbor algorithm. Proceedings of the European Conference of the Prognostics and Health Management Society.
Tikhamarine, Y., Souag-Gamane, D., Ahmed, A.N., Kisi, O., El-Shafie, A., 2020. Improving artificial intelligence models accuracy for monthly streamflow forecasting using grey Wolf optimization (GWO) algorithm. J. Hydrol. 582, 124435.
Tsai, W.P., Chang, F.J., Chang, L.C., Herricks, E.E., 2015. AI techniques for optimizing multi-objective reservoir operation upon human and riverine ecosystem demands. J. Hydrol. 530, 634–644.
Tsai, W.P., Huang, S.P., Cheng, S.T., Shao, K.T., Chang, F.J., 2017. A data-mining framework for exploring the multi-relation between fish species and water quality through self-organizing map. Sci. Total Environ. 579, 474–483.
Wang, W.C., Chau, K.W., Cheng, C.T., Qiu, L., 2009. A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. J. Hydrol. 374 (3–4), 294–306.
WL/Delft Hydraulics, 1995. SOBEK 1.0: User's Guide. Delft, The Netherlands: WL/Delft Hydraulics and Ministry of Transport and Management.
Xie, X.L., Beni, G., 1991. A validity measure for fuzzy clustering. IEEE Trans. Pattern Anal. Mach. Intell. 13 (8), 841–847.
Yamazaki, D., Kanae, S., Kim, H., Oki, T., 2011. A physically based description of floodplain inundation dynamics in a global river routing model. Water Resour. Res. 47 (4) https://doi.org/10.1029/2010WR009726.
Yang, S.Y., Chan, M.H., Chang, C.H., Chang, L.F., 2018. The damage assessment of flood risk transfer effect on surrounding areas arising from the land development in Tainan, Taiwan. Water 10 (4), 473.
Yaseen, Z.M., El-Shafie, A., Jaafar, O., Afan, H.A., Sayl, K.N., 2015. Artificial intelligence based models for stream-flow forecasting: 2000–2015. J. Hydrol. 530, 829–844.
Yoo, C., Hong, S.H., Choi, H., Nam, K.W., 2013. A study on flood area typology using the inundation trace map-a case study of Busan metropolitan city. J. Korean Soc. Survey., Geod., Photogram. Cartogr. 31 (5), 393–400.
Zhang, W., Wang, J., Jin, D., Oreopoulos, L., Zhang, Z., 2018. A deterministic self-organizing map approach and its application on satellite data based cloud type classification. 2018 IEEE International Conference on Big Data (Big Data), December, 2018, IEEE, pp. 2027–2034.
Zhao, G., Xu, Z., Pang, B., Tu, T., Xu, L., Du, L., 2019. An enhanced inundation method for urban flood hazard mapping at the large catchment scale. J. Hydrol. 571, 873–882.