We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS

BOOK CITATION INDEX

INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Machine Learning-Based Method for Urban Lifeline System Resilience Assessment in GIS*

Wenjie Huang and Mengzhi Ling

Abstract

System resilience, the capability of a system to sustain and recover from deliberate attacks, accidents, or naturally occurring threats or incidents, is a key property to measure the degree of robustness and coupling effect of complex system. The systems of waste disposal, urban water supply, and electricity transmission are typical systems with complex and high coupling features. In this chapter, a methodology for measuring the system resilience of such systems is proposed. It is a process of integrated decision-making which contains two aspects: (1) a five-dimensional indicator framework of system resilience which includes attributes in infrastructural, economic, and social sectors and (2) a hybrid K-means algorithm, which combines entropy theory, bootstrapping, and analytic network process. Through utilizing real data, the methodology can assist to identify and classify the level of system resilience for different geographical regions which are sustained by lifeline systems. The calculation of algorithm, visualization of processed data, and classification of resilience level can be finally realized in geographic information system. Through utilizing by regional governments and local communities, the final result can serve to provide guideline for resource allocation and the prevention of huge economic loss in disasters.

Keywords: system resilience, geographic information system (GIS), analytical network process, K-means

1. Introduction

Urban lifeline systems typically consist of critical infrastructure systems such as water, gas, electricity, communication, and transportation [1, 2]. In the modern era, these systems serve to maintain the function and ensure the quality of life for cities and their citizens. With the rapid urbanization and economic development, the demand of cities is increasing, and different regions of urban areas are interconnecting. This draws attentions to government and communities about the performance of lifeline system under the stresses of emergent status in disasters. Systematical methodologies and techniques, thus, are sought by the government and communities for identifying, analyzing, and prioritizing the resilient capacity

---

of cities. The report of the United Nations Office for Disaster Risk Reduction indicated that "$366 billion in direct damages and 29,782 fatalities worldwide" are caused by natural disasters in 2011. This signifies that enhancing the resilient ability of urban lifeline systems can be a way of ensuring massive savings through risk reduction and expeditious recovery. However, there is currently no precise platform to analyze and visualize system resilience to support urban planning; furthermore, there lacks systematical methodologies for assessing multidimensional complex lifeline system resilience. This chapter proposed an integrated decision-making process for evaluating and classifying the system resilience levels of the urban lifeline systems of urban areas. It is a methodology based on weighted multi-criteria indicator data. This methodology can serve as guidelines for governments and communities to effectively reallocate resources for preventing economic loss and to conduct infrastructure reconstruction in those low resilience level areas. Particularly, the contribution of this chapter has four main parts:

1. For evaluating the system resilience systematically, a multi-attribute indicator framework of system resilience is integrated. This framework has been proposed in literature, which includes five-dimensional attributes.

2. For classifying the resilience level of urban regions, a hybrid K-means algorithm based on weight-adjusted data is proposed.

3. For effectuating the algorithm and visualizing the clustering results of resilient level, the GIS platform is introduced.

This chapter is structured by four sections. First, the literature review is embedded in the Introduction. Second, the methodology is elaborated and interpreted. Third, the application of the proposed methodology is illustrated through a sample case of salt tide hazard in China. Fourth, an overview of the contribution and applicability of this study and recommendation for further study are given.

1.1 Literature review

1.1.1 Resilience of complex systems

According to the Presidential Policy Directive (PPD)-21 on Critical Infrastructure Security and Resilience, the terminology "resilience" represents "the ability of a system to prepare for and adapt itself to changing conditions, and rapidly withstand and recover from disruptions." For urban systems, the domain of disruptions includes deliberate attacks, accidents, or naturally occurring threats or incidents [3]. In order to assist our cities to resist in and recover from disasters, many studies are focusing on establishing system resilience assessment. For instance, Francis and Bekera [25] proposed an assessment framework for system resilience which consists of five parts: "system identification, vulnerability analysis, resilience object setting, stakeholder engagement and resilience capacities." They also developed a triangle system resilience structure which includes "three pillars" of resilient capacity: absorptive, adaptive, and recovery ([4], p. 92). According to their definitions, absorptive capacity is the degree at which a system can absorb the impacts of perturbation and minimize consequences with little efforts ([4], p. 94). Adaptive capacity is the ability of a system to adjust to undesirable situations by undergoing minor changes ([4], p. 94). Recovery capacity is characterized by system reliability and the rapidity of the system returning to normal or improved operations ([4], p. 94).

City is an organism that consists of the physical environment and human beings. The interactions between the environment and human manifest the complexity of urban systems, that is, the mutual coupling among the environmental system and economic system, social systems, and other systems. Classical studies of system resilience, however, fail to take this complexity and attendant uncertainty of urban system into account [4]. In addition to evaluating the quality of lifeline infrastructure and the diversities in socioeconomic and cultural characteristics, the interactions between physical and nonphysical systems can also significantly contribute to the resilience of urban systems [5, 6].

Through studying 100 cities worldwide, an international construction firm, Arup, proposed the City Resilience Index (CRI) [7]. CRI is an urban resilience assessment framework structured by triple-level which consists of 4 dimensions (e.g., health, well-being, economy, and society), 12 goals (e.g., minimized human vulnerability and sustainable economy), and 52 indicators (e.g., protection of livelihoods following a shock and well-managed public finances) [7].

The assessment of complex system resilience enables decomposing resilience into its individual attributes and then organizing the attributes into an organizational tree of indicators from different aspects and levels. Accordingly, resilience could be evaluated through a combination of these indicators [8]. This provides ideas for collecting data from various sub-systems of urban organism, but a complex methodology concerning the interdependencies of such variables is required [3, 9, 10].

Pregenzer [11] believes that, in future nuclear system resilience research, new methodologies for soliciting expert opinions and analyzing historical data will be required to assess the relative strengths and potential unintended consequences of nonproliferation strategies. In a recent research by [12], for studying the long-lived socio-ecological systems in a causal loop diagram, multi-methods are applied. The methods include observation, expert opinion, and archive research. However, existing resilience assessment framework such as CRI is generally dependent on the method of expert scoring. Possible subjective bias caused by the method may reduce the reliability of the assessment results and have negative influence on further implement.

1.1.2 Resilience evaluation methods

The ability of a system to adequately and efficiently respond to external risks and internal instability is crucial in the context of urbanization and socioeconomic development. It is particularly critical to integrated information and communication systems and factors [9, 13]. For urban infrastructure systems, research has been done on developing generalized models to analyze the interdependencies among systems. One of the methods is called multilayer infrastructure network framework where economic flows and information flows transmit between different system layers. Accordingly, the system interdependency could be derived by computable general equilibrium (CGE) theory [10]. Mathematical modeling dealing with economic impact and recovery time is most commonly applied for resilience of supply chain system and related infrastructure. In make-to-stock systems, integral of the time absolute error is an appropriate control engineering measure of resilience for system inventory level and shipment rates [14]. Static assessment of system resilience helps to reflect intrinsic properties of certain lifeline systems. The work [15] proposes a qualitative approach to measure the resilience of residential buildings in various exogenous hazards scenarios, by using different parameters in the loss function and recovery function. Simulation techniques such as Monte Carlo analysis are used to conduct numerical studies on resilience evaluation and estimation [15].

1.1.3 Analytical network process

In the areas of risk assessment and decision analysis, analytical network process (ANP) or analytical hierarchy process (AHP) is widely used to assess the key factors of risks and analyze the impacts and preferences of decision alternatives [16]. The work [17] realizes the dynamic analysis of interaction among five urban elements: nature, society, economy, technology, and management and evaluation of regional flood hazard resilience through adopting ANP. However, subjective judgments of personnel and the scorings of decision-makers are common limitations of ANP and AHP. The ordered weighted average approach is a countermeasure to the constraints. It assists to eliminate these bias and subjective through making a series of local aggregations at each level of AHP [18, 19]. More advanced theories such as stochastic AHP transform the preference of decision-makers from stochastic pairwise comparisons to certain probability distributions with minimal information loss and gain optimal strategies by solving the problem of nonlinear programming model [20]. Other indicator weighting methods include robustly and stochastically weighted multi-objective optimization model [21], interval judgment matrix [22], and stochastic dominance [23].

Based on literature review, there exist several shortcomings in current system resilience evaluation methodologies. First, the mathematical definition of "system resilience" is absent which causes the constraint to measure system resilience quantitatively and elaborately. It means that in the study of system resilience assessment, instead of relying on a fixed mathematical model, new techniques purely based on real data are required. Second, there lacks a method to effectively integrate the human knowledge of decision-makers and experts into the process of system resilience evaluation. Accordingly, the methodology proposed in this chapter aims to improve the evaluation of system resilience in three aspects: (1) formulating an indicator pool that integrated factors in existing resilience assessment, (2) involving human knowledge to the process through inviting evaluators to select proper indicators from the pool for specific scenario of lifeline system and disaster (e.g., the water supply system facing salt tide and the electricity generation system experiencing hurricane) and to score, and (3) introducing techniques of machine learning to reduce the bias effects of human interpretations.

2. Methodologies

This section elaborates the proposed methodology for system resilience assessment of lifeline system. The methodology aims to identify and classify the resilience levels of urban regions supported by the lifeline system. For instance, it can be used to evaluate the resilience levels of residential areas, commercial areas, and industrial parks supported by the urban water supply system. This methodology has five parts:

1. An integrated system resilience indicator framework: [24] collected 36 system resilience assessment frameworks and proposed five dimensions of resilience assessment—material and environmental resources (M&ER), society and well-being (S&WB), economy (E), built engagement and infrastructure (BE&I), and governance and institutions (G&I) [24]. The proposed study adapted result in [24] as a resilience evaluation framework and called it MSEBG framework.

2. Adopting three pillars of system resilience capacity: the absorptive, adaptive, and recovery capacities from [25]. The evaluators, that is, the experts, then

chose the appropriate indicators from the above five dimensions under each pillar according to different hazards (e.g., the experts consider that "diverse population composition" is significant for the recovery of water supply systems during flood; therefore, they will classify this indicator under the category of "recovery capacity").

3. Using ANP model to assign the weights of the indicators. Explicitly, the weights represent the relative importance of each indicator from decision-maker's perspective.

4. Developing a hybrid K-means algorithm for adjusting weight and clustering: the algorithm proposed in this chapter combines standard K-means algorithm with the entropy theory and bootstrapping method. The algorithm uses real data of the indicator as input and then identifies and classifies the resilience level of different regions supported by the lifeline system.

5. Building GIS for calculation and visualization of resilience clustering results.

In **Figure 1**, the connection between various methods applied in this study is shown. In this methodology, the indicator framework will be built according to the selected appropriate indicators for measuring the resilience of specific system, and the raw data of indicators will be assigned according to the "three pillars" of capacities. Decision-makers use ANP model to generate their weights. The raw data and the weights are the inputs of the hybrid K-means algorithm; the algorithm contributes to cluster all the communities into groups with different resilience levels. These result data are finally transferred into a GIS platform; and the visual results can be shown.
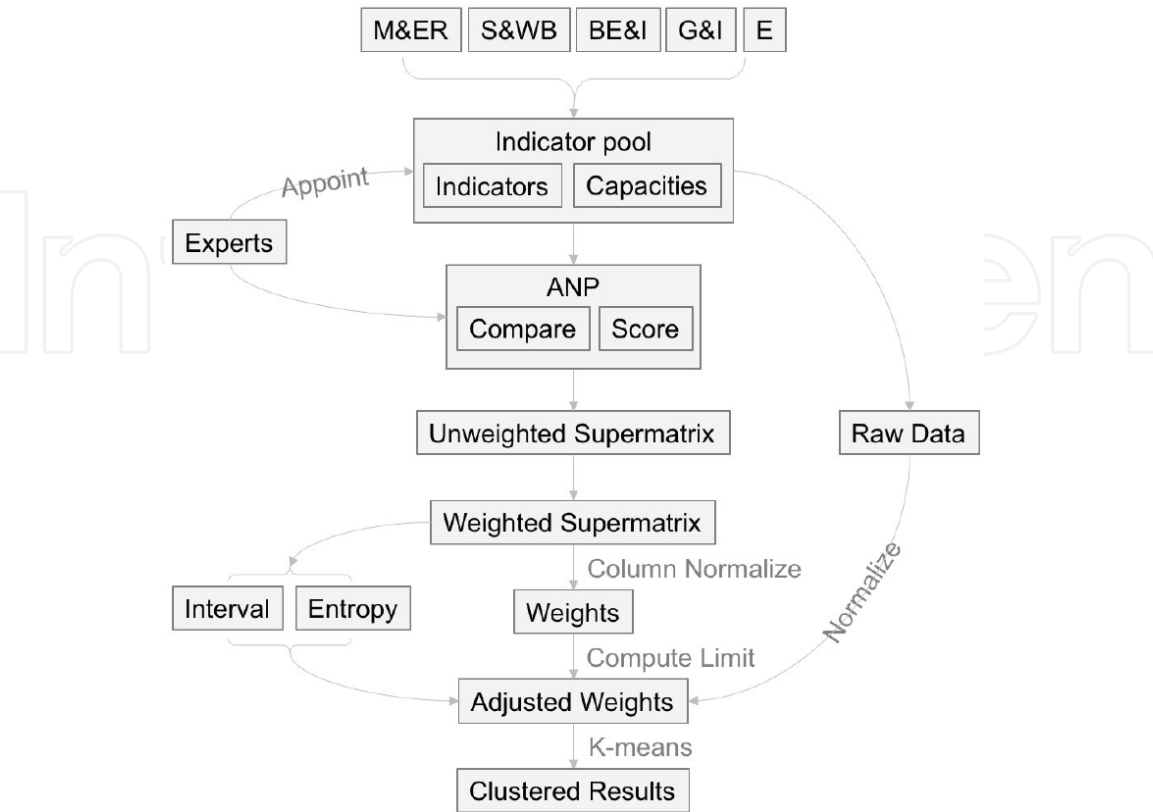

Figure 1.
Decision-making flow.

2.1 System resilience evaluation indicator framework

In general existing urban resilience assessment frameworks consist of different dimensions of urban system and various indicators for evaluation. In this chapter, the indicator framework directly applies the indicator matrix proposed by [24], who collected attributes of urban resilience from 36 assessment frameworks and categorized them into 5 dimensions with 122 indicators. Subsequently, indicators from these dimensions are classified into the "three pillars" of system resilience by evaluators. This will guide decision-makers to identify and extract key indicators for resilience assessment of the complex lifeline systems.

After construction of framework, the experts can select the detailed indicators from the indicator pool proposed by [24], classify the indicators into "three pillars" of capacities, and then derive the indicators' interdependence. A two-layer ANP model is constructed based on these indicators, in which the top layer is the "three pillars" of capacities and the bottom layer is constructed by the chosen indicators.

The experts choose an integer from −10 to 10 to represent the relative importance of any two indicators according to their perspective (for comparing "protection of wetlands and watersheds" and "availability and accessibility of resources," negative value means that "protection of wetlands and watersheds" is more important than the latter; and positive value means that "availability and accessibility of resources" are more important than the previous). The absolute value reveals the relative importance of an indicator inside a pair, i.e., selecting −10 and 10 means that one indicator strongly dominates and is dominated by the other, respectively, while 0 means they are almost equally important.

The ANP algorithm strictly refers to decision-making process introduced by [26]. The ANP will then generate weighted and unweighted super-matrices from the expert scoring. The row and column number of unweighted super-matrices equals to the number of indicators. The weighted super-matrices store the weights of indicators under the same capacity category and the weights between capacities. Explicitly, the weighted super-matrices are computed by the column normalization of unweighted super-matrices. The weight for all the indicators equals to the limited matrix iteratively computed and convergent from weighted super-matrices.

2.2 Hybrid K-means algorithm

The hybrid K-means algorithm is developed in this section: the overview of the algorithm is presented first. The interpretations at each step of this algorithm are presented later. The contents for the algorithm are shown as follows:

- N: the total number of all regions

- R: the total number of indicators

- $D \in \mathbb{R}^{N \times R}$: the column-normalized indicator data with N regions and R indicators. Column normalized means that each original element in the column is divided by the maximum value in that column. The calculation enforces each element in the matrix D taking value inside $[0, 1]$ and thus normalizes all the data into the same scale. Normalization is a standard procedure before performing K-means algorithm, when the indicators are of incomparable physical units.

- $Ws \in \mathbb{R}^{R \times 1}$: the weight of each indicator by the ANP model

- $SR \times R$: the initial super-matrix generated by the ANP model with R indicators

- a_{ij}, $i, j \in \{1, ..., R\}$: each element in the matrix S for a given super-matrix S

- $H \in \mathbb{R}^{1 \times R}$ and $L_i, U_i \in \mathbb{R}^{R \times 1}$, $i \in \{1, ..., R\}$: entropy values

- $[L_i, U_i]$, $i \in \{1, ..., R\}$: closed intervals

The contents of hybrid K-means algorithm are summarized in Algorithm 1.

In Algorithm 1, Step 1a computes a confidence interval for each indicator based on bootstrapping method. Step 1b computes the entropy value from the super-matrix S. In Step 2, the optimal weight $\hat{W}o$ equals to the optimal solution of the optimization problems (1)–(4), where the weight confidence interval constraint is incorporated as (3). In (4), to ensure that each weight is greater than zero, we set $\epsilon = 0.001$. In Step 3, Ws represents the subjective weight, and Wo represents the objective weight. Step 4 outputs the weight-adjusted indicator data. Finally, Step 5 implements the classical K-means algorithm. The structure of K-means algorithm follows what have been proposed in [27]. Set the number of clustering groups as k. After implementing Algorithm 1 and getting the clustering results, the term m_j, $j = 1, 2, ..., k$ is used to denote the centroid of each clustering group. The norm of each cluster centroid is computed as $C_j = \left\| m_j^{(T)} \right\|_2$, $j = 1, 2, ..., k$ and is then ranked from low to high, that is, $C_{(1)} \leq C_{(2)} \leq ... C_{(k)}$. Identifying the system resilience level of points in clustering group is based on the norm of the clustering centroid. The resilience level of the group with the centroid $D_{(j)}$ equals j.

2.2.1 Entropy and bootstrapping

In statistics, bootstrapping methods refer to the tests or metrics that rely on random sampling with replacement. Bootstrapping method allows assigning measures of accuracy (bias, variance, confidence intervals, prediction error, etc.) to sample estimates [28, 29]. In this chapter, Algorithm 2 shows the bootstrapping method that outputs the confidence interval for the weight of each indicator depending on super-matrix S in ANP and thus provides the feasible interval for the weight of each indicator for further adjustment and optimization.

The feasible interval of each indicator i is computed by resampling n samples from the initial sample space $\{a_{i1}, a_{i2}, ..., a_{iR}\}$, and then repeat such procedure for B times. Then compute the average value of the samples in the resampled space for each $b = 1, 2, ..., B$. Step. 2 and 3 in Algorithm 2 rank the average value of resample batches and then compute the estimated value for the $1-\alpha$ confidence interval of the average value. The feasible interval of each indicator equals the estimated confidence interval, i.e., $[L_i, U_i] = \left[\hat{\theta}_i^{*, (k1)}, \hat{\theta}_i^{*, (k2)} \right]$. Further, the optimized models (1)–(4) with entropy objective are interpreted. Entropy theory has been applied to wide range of system resilience assessments from engineering and economics to anthropology and social ones [30–33]. Entropy indicates the degree of disorder, uncertainty, or lack of information about the configuration of system modules [34]. The lower the entropy value, the higher is the information utility it has. Here, the entropy value H computed by Algorithm 3 follows the definition in [35–37] and evaluates the information utility and reliability for the weights of indicators generated by ANP.

Algorithm 1. Hybrid K-means.

Input: Original indicators data $D \in \mathbb{R}^{N \times R}$; Weight matrix by ANP model $Ws \in \mathbb{R}^{R \times 1}$, and Super matrix $S \in \mathbb{R}^{R \times R}$;

Step 1: for $i, j = 1,...,r$ **do.**

Step 1a: Compute the bounds $[L_i, U_i]$, $\in \{1,...,R\}$ by row data $\{a_{i1}, a_{i2}, ..., a_{iR}\}$ $i = 1,...,R$ from S using bootstrapping methods, summarized in *Algorithm 2*;

Step 1b: Compute the Entropy value H, based on column data $a_{1j}, a_{2j}, ..., a_{Rj}$ from S, summarized in *Algorithm 3*;

Step 2: Define $Wo \in \mathbb{R}^{R \times 1}$ with each element as w_i, $i \in \{1,...,R\}$, given a positive threshold $\epsilon > 0$, and solve the linear program

$$\min_{Wo} \quad H \cdot Wo \tag{1}$$

$$\text{s.t.} \quad \sum_{j=1}^{R} w_i = 1, \tag{2}$$

$$L_i \leq w_i \leq U_i, \qquad\qquad i = 1, ..., R, \tag{3}$$

$$w_i \geq \epsilon, \qquad\qquad i = 1, ..., R, \tag{4}$$

which returns the optimal $\hat{W}o = \{\hat{w}_i\}$, $i = 1,2,...,R$.

Step 3: Denote $Ws_{(r \times 1)} = \{w'_i\}$, $i = 1,2,...,R$. Compute the adjusted weights by

$$W = \left\{ \frac{\hat{w}_i \times w'_i}{\sum_{i=1}^{R}(\hat{w}_i \times w'_i)} \right\}, i = 1, ..., R, \tag{5}$$

Step 4 : *Construct weight matrix* $W' \in \mathbb{R}^{N \times R}$ *by N copies of array* $W^{\top}$. *Compute the weight*

adjusted indicator data by Hadamard product of weight matrix W' and data matrix D, i.e., $D' = W' \circ D$

Step 5: Implement K-means algorithm on D'

Return clustering result on N regions;

Algorithm 2. Bootstrapping

Input: Define an initial sample space $\{a_{i1}, a_{i2}, ..., a_{iR}\}$ for each $i \in \{1, 2, ..., R\}$ from the Weighed supermatrix S; Confidence level α;

Step 1: Randomly select B batches of samples with size n from initial sample space, for each, $i \in \{1, 2, ...R\}$,

$$A_i^{*, b} = \left(a_{i1}^{*, b}, a_{i2}^{*, b}, ..., a_{in}^{*, b} \right), b = 1, 2, ..., B,$$

for $b = 1,2,...,B$ **do**, compute $\hat{\theta}_i^{*, b} = \frac{1}{n} \sum_{j=1}^{n} a_{ij}^{*, b}$, for $i \in \{1, 2, ...R\}$, and $b = 1,2,...,B$.

Step 2: Sort the value of $\hat{\theta}_i^{*, 1}, \hat{\theta}_i^{*, 2}, ..., \hat{\theta}_i^{*, B}$ from low to high, and obtain $\hat{\theta}_i^{*, (1)} \leq \hat{\theta}_i^{*, (2)} \leq ... \leq \hat{\theta}_i^{*, (B)}$, for each $i \in \{1, 2, ...R\}$.

Step 3: Set $k_1 = \alpha \times B/2$, and $k_2 = B - \alpha \times B/2$, let $\hat{\theta}_i^{*, (k_1)}, \hat{\theta}_i^{*, (k_2)}$ be the estimation of $\hat{\theta}_i^{*, \alpha/2}, \hat{\theta}_i^{*, 1-\alpha/2}$, that is, the estimation of $1 - \alpha$ confidence interval of statistics θ.

Algorithm 3. Computation of Entropy.

Input: Super matrix $S = \{a_{ij}\} i = 1,...,R, j = 1,...,R$.

Step 1: Normalize the matrix element by $a'_{ij} = \frac{a_{ij}}{\sum_{j=1}^{R} a_{ij}}$, $i, j \in \{1, ..., R\}$,

Step 2: Normalize by column to obtain $f_{ij} = \frac{a'_{ij}}{\sum_{i=1}^{R} a'_{ij}}$.

Step 3: for $j = 1, \ldots R$ **do**, compute entropy by definition:

$H_j = -\frac{1}{\ln(r)} \sum_{i=1}^{r} f_{ij} \ln\left(f_{ij}\right)$.

Return Entropy value H.

For S, each column vector $\left(a_{1j}, a_{2j}, \ldots, a_{Rj}\right)^{\top}, j = 1, 2, \ldots, R$ denotes the weighting strategy under the jth indicator's criterion; therefore, the weighted aggregation of each column's output entropy represents the total uncertainty metric of the indicator system. Summarized by optimization models (1)–(4), the algorithm seeks to find the optimal weight strategy $\hat{W}o$, to minimize the weighted aggregation entropy value $H \cdot Wo$, with respect to the feasible interval computed from bootstrapping confidence interval. Namely, the algorithm intends to find an optimal weight strategy to maximize the overall information utility. Such strategy would help to eliminate the side effects of subjective judgments from the experts. Superimposing the subjective and objective weights together realizes a comprehensive weight assignment strategy.

3. Case study

Case study of system resilience evaluation for a water supply system in the risks of salt tide is presented in this section. The case study validates the proposed methodology.

3.1 Background information

This section evaluates the Chenhang reservoir water supply system resilience under the salt tide hazards in the estuary of Yangtze River. The formal definition of salt tide is the emergency situation that chloride concentrations in water body exceed the national standard level (250 mg per liter of water). Salt tide destroys the quality of water, results in soil salinization in coastal areas and cities, and has negative impacts on production and human daily life. Recently, salt tide has already become one of the most internationally concerned disasters of coastal cities. The invasion of saltwater during salt tide will limit the access of high-quality municipal and industrial water and will cause water shortage and scheduling problems in some megacities of China located in estuarine and coastal areas, such as Shanghai. Thus, water shortage problem has become one of the main obstacles that obstructs the construction of eco-city and sustainable development. By experience, the intensity of salt tide intrusion changes with the period of tides, showing its periodic properties. In general, September to April next year is the period of time influenced by salt tide of water intake in Yangtze River. Each intrusion of salt tide lasts for 6–7 days. Since there exist multiple factors that affect the duration and extent of hazards (e.g., Yangtze River hydrology, chase traffic, weather, and wind), it is usually difficult to make detailed and accurate prediction for each intrusion. In recent years, the hazards have become more severe in the following aspects: long intrusion duration, high frequency, short interval time between intrusions, and independence of Yangtze River runoffs. Given the fact that the "Three Gorges Project" increased the ability to implement different water strategies at the upstream of Yangtze River, the extreme hydrological hazards occurred more frequently, which make the research on how salt tide influences the water supply system more practical and crucial [38–40].

According to chemistry knowledge, physical and chemical contaminates such as chloride concentration, ammonia, oxygen consumption, and total iron play a role as indicators to show the degree of salt tide level. These indicators reach peak values during December to February [41]. Accordingly, the whole period can be categorized into three different periods:

- Period I: September to December, when the chloride concentration in water accumulates up to 250 mg/L

- Period II: December to February (next year), when the chloride concentration abruptly reaches approximately the peak value (250–450 mg/L)

- Period III: From February to April, when salt tide recedes and chloride concentration returns to normal level

The goal of this case study is to study the system capacities in terms of preparation for salt tide hazards, facing salt tide hazards, and recovering from salt tide hazards. Period I is the preparation period, when the salt tides gradually affect the water supply system. In Period III, the recovery phase would occur, after the severe disturbances of salt tide. Intuitively, for Period I, decision-makers should focus more on increasing the adaptive capacity of system when prevention and emergency are crucial to system against increasing salinity. In Period II, absorptive capacity should be emphasized, since the system needs to absorb perturbations and minimizes the consequence when contaminants immediately reach the peak. In Period III, recovery capacity is more important indicator when water supply system needs redesigning and rebuilding. Thus, for each period, the decision-making process can be implemented dynamically, and the weight evaluations of indicators and clustering results can be derived step by step.

In the studied area, Chenhang reservoir supplies water to districts (e.g., Baoshan, Jiading, Putuo, Zhabei, and Hongkou) and towns (e.g., Gaodong, Gaoqiao, and Gaoxing inside Pudong New District) in Shanghai. The whole study region is zoned into 29 communities on the basis of administrative divisions shown in **Figure 2**. A unique FID number is assigned to each community.
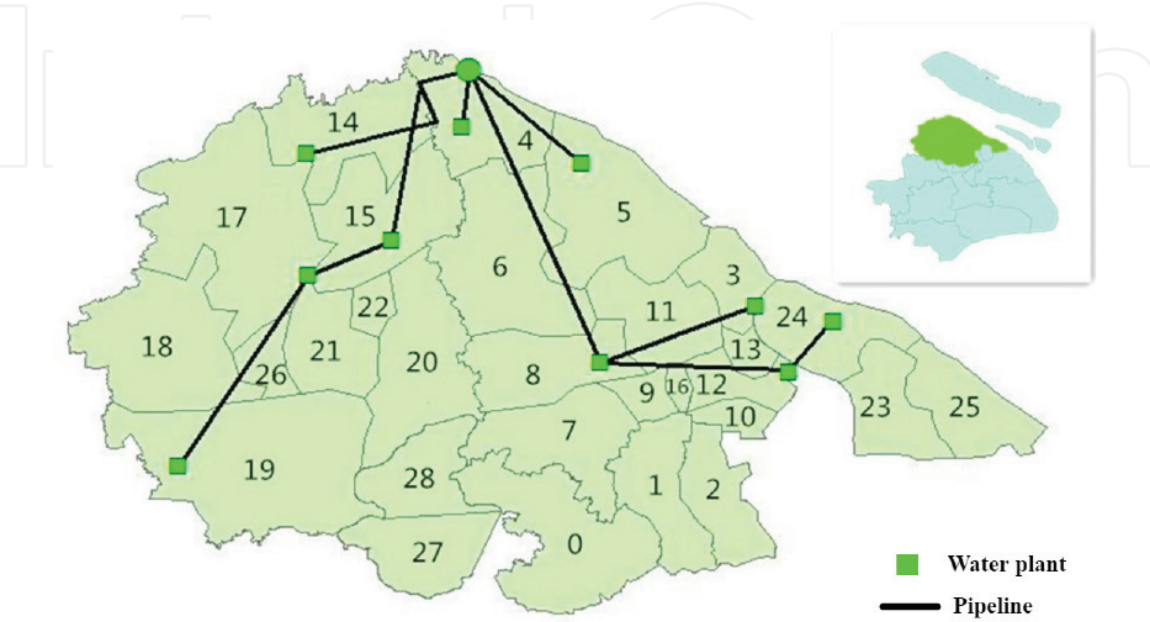

Figure 2.
Basic GIS Layer for Chenhang reservoir water supply system.

3.2 Resilience evaluation results

The resilience evaluation results are presented and analyzed in this subsection.

3.2.1 ANP model

In Section 2, the MSEBG indicator framework for system resilience and ANP model were exhaustively introduced. The indicators selected from the experts for this case study are shown in **Table 1**.

In this case study, 50 experts were involved. Those experts are chosen based on the selection method proposed in [42]. Half of the experts are professional technicians from lifeline system industry or experts in the field of system resilience, and another half of the experts are part-time college students majoring in Environmental Engineering and Urban Design. With solid technical and practical backgrounds,

	Absorptive capacity
M & E R	Protection of wetlands and watersheds
	Availability and accessibility of resources (water)
S & WB	Degree of connectedness across community groups
E	Investment in green jobs and green economy
BE & I	Robustness and fortification
	Spatial distribution of critical infrastructure
	Regular monitoring, maintenance, and upgrade of critical infrastructure
G & I	Decentralized responsibilities and resources
	Understanding risk patterns and trends
	Adaptive capacity
M & E R	Material and resource management
S & WB	Place attachment and sense of community and pride
E	Income and poverty
	Stability of prices and incomes and property value
BE & I	Redundancy of critical infrastructure, facilities, and stocks
	Consolidation of critical utilities and collaboration between utility providers
	Diverse and reliable information and communication technology (ICT) networks
G & I	Strong leadership
	Drills and exercises
	Recovery capacity
M & E R	Ecosystem monitoring and protection
	Protection of wetlands and watersheds
S & WB	Population composition
	Past experience with disaster recovery
E	Insurance and social welfare
G & I	Efficient management of resources

Table 1.
Indicators for water supply system.

they can understand and evaluate system resilience and provide scores for the corresponding indicators.

The ANP model for this case study is presented in **Figure 3**. The connection arc shows the independent relationship between two indicators.

3.2.2 Results and analysis for clustering

The data of the selected indicators (denoted by *D* in Algorithm 1) comes from *2011 Statistical Yearbook of Shanghai* issued by Shanghai Municipal Statistics Bureau (SMSB) and 2011 statistical yearbooks issued by Statistic Bureau of Shanghai. The corresponding semantic-type or non-quantifiable indicator data are transformed into real number value by expert scoring method. As summarized in Section 2, the input data *D* is preprocessed with column normalization to eliminate the bias effects.

Figure 4 illustrates the system resilience levels of all communities in the region under three periods. Particularly, all the districts are clustered into four resilience levels: Level 1 (particularly low), Level 2 (relatively low), Level 3 (relatively high),
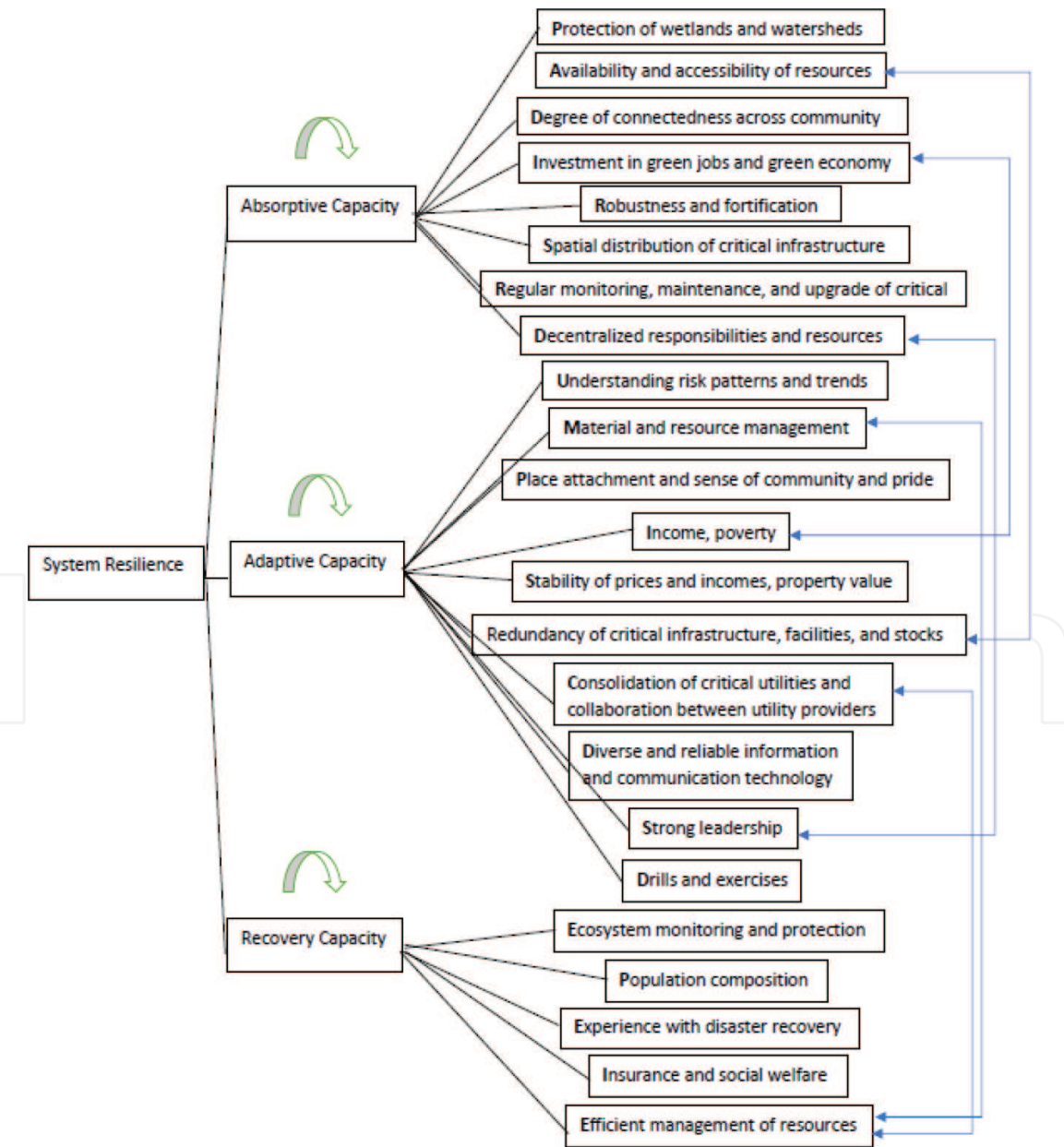

Figure 3.
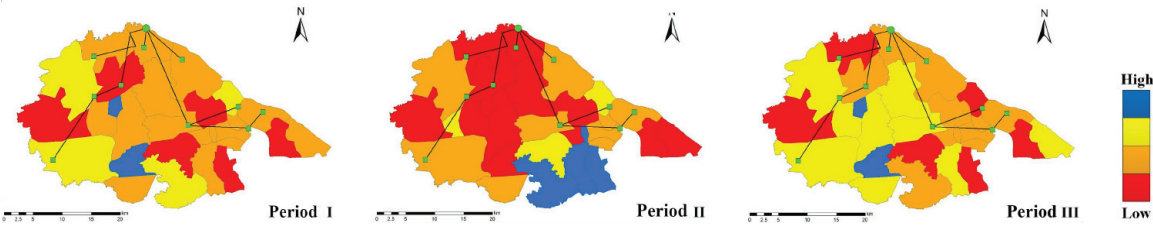*ANP model for water supply system.*

Figure 4.
System resilience level under Periods I, II, and III.

and Level 4 (high). The four levels are represented by colors red, orange, yellow, and blue. Such four level clustering concepts are inherited from warning signals of meteorological disasters on [43]. The visualized clustering result from GIS system enables decision-makers to prioritize crucial prevention and protections to the areas with low system resilience level and act dynamically when water supply systems face different periods of salt tide. **Table 2** shows that the resilience levels of most of the communities are clustered into Level 1 and Level 2. **Figure 4** shows a

	Period I		Period II		Period III	
j	D_j	**Region unit FID**	D_j	**Region unit FID**	D_j	**Region unit FID**
1	0.8990	2, 7, 11, 15, 18, 23, 26	0.9353	4, 6, 9, 11, 14, 15, 16, 18, 20, 21, 22, 23, 25	0.9083	2, 7, 11, 14, 18, 23, 26
2	1.0175	1, 2, 4, 5, 6, 8, 9, 10, 12, 13, 14, 16, 20, 21, 24, 25, 27	1.2046	5, 8, 10, 12, 13, 17, 19, 24, 27	1.0811	1, 4, 5, 6, 7, 8, 9, 10, 12, 13, 15, 16, 20, 21, 24, 25, 27
3	1.2638	0, 3, 17, 19	1.3689	3, 7, 26	1.2039	0, 3, 17, 19
4	1.7721	22, 28	1.8976	0, 1, 2	1.1610	22, 28

Table 2.
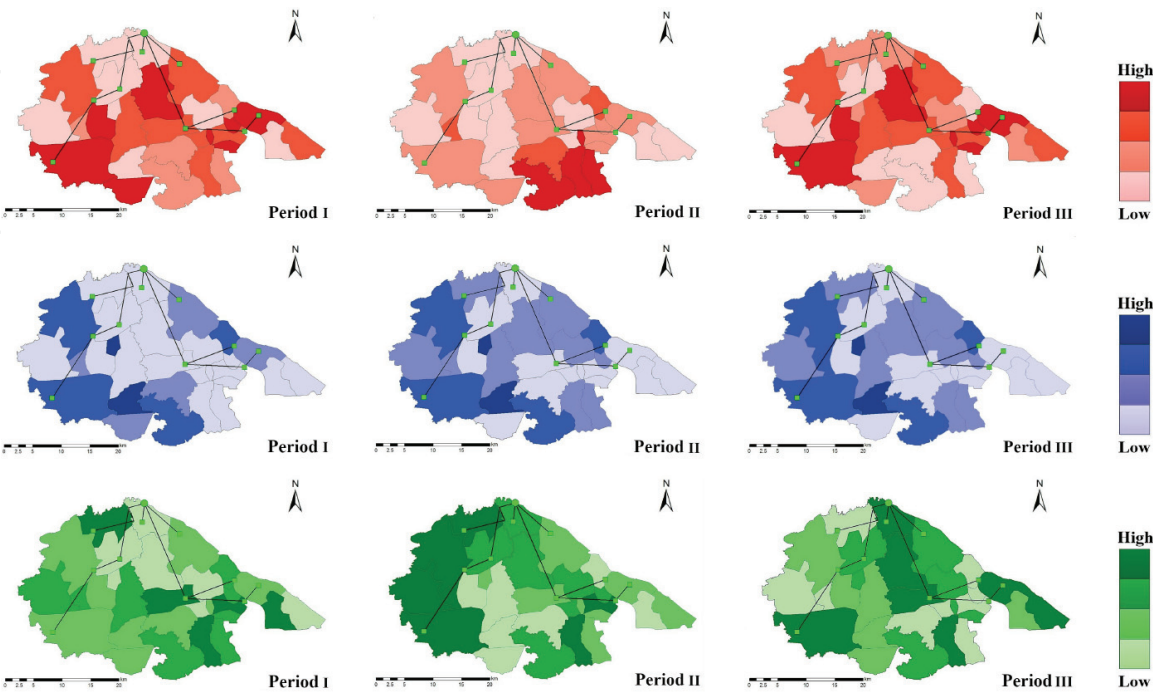*Overall system resilience level analysis results under Periods I, II, and III.*

Figure 5.
Absorptive (red), adaptive (blue), and recovery (green) capacities under Periods I, II, and III.

phenomenon that the communities closer to the reservoir and contained inside the network of the water supply systems have lower resilience levels. These areas show closer exposure to the hazard and have less time to provide rapid responses, and also these communities take charge of the protection of network in disaster; reversely, the communities located far away from the reservoir, as well as the downtown area, have better economic status and higher service level. Absorptive, adaptive, and recovery capacities are analyzed, respectively, for Periods I, II, and III, with the results expressed in **Figure 5**. The complete clustering results are presented in **Table 3** with associated FID. **Figure 5** shows that for different periods, the resilience level of each community varies between three different capacities.

Period I				Period II		Period III	
Capacity		D	Region unit FID	D_j	Region unit FID	D_j	Region unit FID
Absorptive	1	0.6163	4, 11, 14, 15, 18, 22, 23, 28	0.9267	4, 9, 10, 11, 14, 15, 18, 20, 21, 22, 23, 25, 28	0.2692	4, 11, 14, 15, 18, 22, 23, 28, 26
	2	0.7354	0, 2, 3, 7, 26	1.1995	5, 6, 8, 10, 12, 13, 17,19,24,27	0.4120	0, 2, 3, 7
	3	0.7988	1, 5, 8, 9, 12, 17, 20, 25	1.3633	3, 7, 26	0.4597	1,5,8,9,12,17,20,25
	4	0.9418	2, 3, 6, 10, 13, 16, 19, 21, 24, 27	1.8941	0, 1, 2, 16	0.5284	2, 3, 6, 10, 13, 16, 19, 21, 24, 27
Adaptive	1	0.4996	1, 2, 4, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 20, 21, 23, 25	0.0669	4, 7, 8, 9, 10, 12, 15, 16, 21, 23, 24, 25, 27	0.1216	4, 7, 8, 9, 10, 12, 15, 16, 21, 23, 24, 25, 27
	2	0.7357	5, 7, 24, 26, 27	0.0894	1, 2, 5, 6, 11, 13, 14, 18, 20, 26	0.1693	1, 2, 5, 6, 11, 13, 14, 18, 20
	3	0.8538	0, 3, 17, 19	0.1589	0, 3, 17, 19	0.3134	0, 3, 17, 19, 26
	4	1.6285	22, 28	0.3322	22, 28	0.6941	22, 28
Recovery	1	0.3043	4, 6, 9, 11, 15, 22, 25	0.0529	3, 7, 20, 21, 23, 27, 28	0.8393	2, 3, 7, 9, 12, 13, 14, 18, 26
	2	0.3119	5, 7, 13, 16, 17, 19, 20, 21, 24, 27, 28	0.0573	2, 5, 8, 10, 16, 22, 24, 25	0.9826	17, 20, 21, 23, 27, 28
	3	0.3295	0, 2, 3, 10, 13, 18, 20, 26	0.0607	0, 4, 6, 9, 11, 13, 15	0.9858	0, 5, 10, 11, 15, 16, 22
	4	0.3460	1, 8, 12, 14, 23	0.0667	1, 12, 14, 17, 18, 19, 26	1.1447	1, 4, 6, 8, 19, 24, 25

Table 3.
Three capacity analyses under Periods I, II, and III.

4. Conclusion

The objective of this study was to propose a decision-making process and a methodology of system resilience assessment for urban lifeline systems. In this work, formation of the concept "system" should not only be limited to system infrastructures but also be expanded to the combination of other related complex systems such as demographic system, economic system, and environmental system in the study region. The advantages of using ANP as weight methods for the

indicators come from two aspects: first, ANP structures the decision-making process by considering both the hierarchy relationship and the interdependence between bottom level indicators; and that is what "networks" in "ANP" comes from; and second, ANP generates weights through sequential pairwise comparisons of experts after selecting any two of the indicators, which is a very straightforward approach for real-life application. There exist four advantages for using hybrid K-means algorithm: (1) the number of clustering groups can be set before running the whole process; (2) the performance of K-means algorithm in high-dimensional clustering problems is relatively superior than other clustering algorithms such as fuzzy C-means, mountain, subtractive, hierarchical, and density-based clusterings in terms of quality, accuracy, and computation time [44–46]; (3) it provides the information of central points of each clustering class, which enables decision-maker to compute their distance from the origin and conduct further spatial analysis and (4) it is a machine learning technique, rather than by a formal mathematical metric of "resilience". Thus this model-free method can be implemented without an explicit mathematical definition of "resilience".

Author details

Wenjie Huang[1]* and Mengzhi Ling[2]

1 Department of Industrial Systems Engineering and Management, National University of Singapore, Singapore

2 Department of Architecture, School of Design and Environment, National University of Singapore, Singapore

*Address all correspondence to: wenjie_huang@u.nus.edu

IntechOpen

References

[1] Moteff J, Copeland C, Fischer J. Critical infrastructures: What makes an infrastructure critical? Library of Congress Washington DC Congressional Research Service; 2003

[2] Zhao X, Cai H, Chen Z, Gong H, Feng Q. Assessing urban lifeline systems immediately after seismic disaster based on emergency resilience. Structure and Infrastructure Engineering. 2016;**12**(12): 1634-1649

[3] Ayyub BM. Systems resilience for multihazard environments: Definition, metrics, and valuation for decision making. Risk Analysis. 2014;**34**(2): 340-355

[4] Fiksel J. Designing resilient, sustainable systems. Environmental Science & Technology. 2003;**37**(23): 5330-5339

[5] Horne AL, Haynes RW. Developing measures of socioeconomic resiliency in the interior Columbia River basin; 1999

[6] Official Website of the World Bank-Data-Indicators; http://data.worldbank.org/indicator

[7] RF TRF, Arup. City resilience index: Understanding and measuring city resilience; 2017. http://www.arup.com/cityresilienceindex

[8] Renschler C, Frazier A, Arendt L, Cimellaro GP, Reinhorn A, Bruneau M. Developing the 'PEOPLES' resilience framework for defining and measuring disaster resilience at the community scale. In: Proceedings of the 9th US National and 10th Canadian Conference on Earthquake Engineering (9USN/10CCEE); Toronto. 2010. pp. 25-29

[9] Helbing D. Globally networked risks and how to respond. Nature. 2013; **497**(7447):51-59

[10] Zhang P, Peeta S. A generalized modeling framework to analyze interdependencies among infrastructure systems. Transportation Research Part B: Methodological. 2011;**45**(3):553-579

[11] Pregenzer A. Systems Resilience: A New Analytical Framework for Nuclear Nonproliferation. Vol. 8. Albuquerque, New México: Sandia National Laboratories; 2011. pp. 1-21

[12] Tenza A, Pérez I, Martínez-Fernández J, Giménez A. Understanding the decline and resilience loss of a long-lived social-ecological system: Insights from system dynamics. Ecology and Society. 2017;**22**(2):15

[13] Barrett C, Beckman R, Channakeshava K, Huang F, Kumar VA, Marathe A, et al. Cascading failures in multiple infrastructures: From transportation to communication network. In: 2010 5th International Conference on Critical infrastructure (CRIS); IEEE; 2010. pp. 1-8

[14] Spiegler VL, Naim MM, Wikner J. A control engineering approach to the assessment of supply chain resilience. International Journal of Production Research. 2012;**50**(21):6162-6187

[15] Tokgoz BE, Gheorghe AV. Resilience quantification and its application to a residential building subject to hurricane winds. International Journal of Disaster Risk Science. 2013;**4**(3):105-114

[16] Ergu D, Kou G, Shi Y, Shi Y. Analytic network process in risk assessment and decision analysis. Computers and Operations Research. 2014;**42**:58-74

[17] Sun H, Cheng X, Dai M. Regional flood disaster resilience evaluation based on analytic network process: A

case study of the Chaohu Lake Basin, Anhui Province, China. Natural Hazards. 2016;**82**(1):39-58

[18] Wang YM, Luo Y, Hua Z. Aggregating preference rankings using OWA operator weights. Information Sciences. 2007;**177**(16):3356-3363

[19] Al-Yahyai S, Charabi Y, Gastli A, Al-Badi A. Wind farm land suitability indexing using multi-criteria analysis. Renewable Energy. 2012;**44**: 80-87

[20] Jalao ER, Wu T, Shunk D. A stochastic AHP decision making methodology for imprecise preferences. Information Sciences. 2014;**270**:192-203

[21] Hu J, Mehrotra S. Robust and stochastically weighted multiobjective optimization models and reformulations. Operations Research. 2012;**60**(4):936-953

[22] Conde E, MdlPR P. A linear optimization problem to derive relative weights using an interval judgement matrix. European Journal of Operational Research. 2010;**201**(2):537-544

[23] Yager RR, Alajlan N. Probability weighted means as surrogates for stochastic dominance in decision making. Knowledge-Based Systems. 2014;**66**:92-98

[24] Yamagata Y, Maruyama H. Urban Resilience. Springer; 2016

[25] Francis R, Bekera B. A metric and frameworks for resilience analysis of engineered and infrastructure systems. Reliability Engineering and System Safety. Switzerland; 2014;**121**: 90-103

[26] Thomas L, Vargas L. Decision Making with the Analytic Network Process. German: Springer; 2006

[27] MacKay D. An example inference task: Clustering. Information Theory, Inference, and Learning Algorithms. 2003;**20**:284-292

[28] Tibshirani EB. An Introduction to the Bootstrap. Monographs on Statistics and Applied Probability. New York, London: Chapman & Hall/CRC; 1993. p. 1

[29] Efron B et al. Second thoughts on the bootstrap. Statistical Science. 2003; **18**(2):135-140

[30] Greco R, Di Nardo A, Santonastaso G. Resilience and entropy as indices of robustness of water distribution networks. Journal of Hydroinformatics. 2012;**14**(3):761-771

[31] Zurlini G, Li BL, Zaccarelli N, Petrosillo I. Spectral entropy, ecological resilience, and adaptive capacity for understanding, evaluating, and managing ecosystem stability and change. Global Change Biology. 2015; **21**(4):1377-1378

[32] Comfort LK, Siciliano MD, Okada A. Resilience, entropy, and efficiency in crisis management: The January 12, 2010, Haiti earthquake. Risk, Hazards & Crisis in Public Policy. 2011;**2**(3):1-25

[33] Comfort LK, Oh N, Ertan G. The dynamics of disaster recovery: Resilience and entropy in hurricane response systems 2005–2008. Public Organization Review. 2009;**9**(4):309

[34] Tamvakis P, Xenidis Y. Comparative evaluation of resilience quantification methods for infrastructure systems. Procedia-Social and Behavioral Sciences. 2013;**74**: 339-348

[35] Bailey KD. Social Entropy Theory. Albany, New York: State University of New York Press; 1990

[36] Bailey KD. Entropy systems theory. In: Systems Science and Cybernetics, in Encyclopedia of Life Support Systems (EOLSS), Developed under the Auspices of the UNESCO. Oxford, UK: Eolss Publishers; 2001. [http://www.eolss.net]

[37] Shannon CE. A mathematical theory of communication. ACM SIGMOBILE Mobile Computing and Communications Review. 2001;**5**(1): 3-55

[38] Vineis P, Chan Q, Khan A. Climate change impacts on water salinity and health. Journal of Epidemiology and Global Health. 2011;**1**(1):5-10

[39] Hongyi W. Shanghai drinking water hit by salt tide. ChinaDaily: Society. 2014. http://usa.chinadaily.com.cn/china/2014-02/24/content_17300378.htm [February 24, 2014]

[40] Jiayun K. Salt Tide Worst on Record, Hits 2m. ShanghaiDaily.com: Metro. 2014. http://www.shanghaidaily.com/Metro/environment/Salt-tide-worst-on-record-hits-2m/shdaily.shtml [February 24, 2014]

[41] Li Y, Chen C, Zhang X, Dai J, Xu H. Impact of salt-tide on urban water supply and its countermeasure. Water Resources and Hydropower Engineering. 2008;**10**:005

[42] Miao R, Huang W, Pei D, Gu X, Li Z, Zhang J, et al. Research on lease and sale of electric vehicles based on value engineering. International Journal of Production Research. 2016;**54**(18): 5361-5380

[43] Zheng G. 16th CMA Order. China Meteorological Administration. 2007. http://www.cma.gov.cn/2011zwxx/2011zflfg/2011zgfxwj/201208/t20120803_180761.html [June 11, 2007]

[44] Hammouda K, Karray F. A Comparative Study of Data Clustering Techniques. Ontario, Canada: University of Waterloo; 2000

[45] Chaudhari B, Parikh M. A comparative study of clustering algorithms using WEKA tools. International Journal of Application or Innovation in Engineering & Management (IJAIEM). 2012;**1**(2): 154-158

[46] Ghosh S, Dubey SK. Comparative analysis of k-means and fuzzy c-means algorithms. International Journal of Advanced Computer Science and Applications. 2013;**4**(4):35-39