



COMPREHENSIVE ASSESSMENT MODEL ON ACCIDENT SITUATIONS OF THE CONSTRUCTION INDUSTRY IN CHINA: A MACRO-LEVEL PERSPECTIVE

Bo SHAO¹, Zhigen HU^{1*}, Liyang TONG², Xiazhong ZHENG², Dawei LIU¹

¹State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan, China

²College of Hydraulic & Environmental Engineering, China Three Gorges University, Yichang, China

Received 20 May 2019; accepted 31 July 2019

Abstract. As one of the most high-risk sections, the construction industry has traditionally been the research hotspot. Yet little attention has been paid to macro-level accident situations of the entire industry. Therefore, this study develops a comprehensive assessment model on accident situations of Chinese building industry, aiming at assisting the government to better understand and improve accident situations of the entire industry. Based on China conditions, six indicators related to accident situations are firstly selected to establish an indicator system; then structure entropy weight method is proposed to determine indicator weights, with dynamic classification method to explore the characteristics of accident situations. The results demonstrate that the provinces with poor accident situations account for 53% of all provinces, and they are mainly distributed in the central and western regions of China where there exist the underdeveloped economy. Meanwhile, some provinces experience poor accident situations that could be out-of-control, especially for Hebei. Provinces in the south-eastern and northeastern regions of China perform relatively well, but they still have much improvement room for accident situations. The findings validate the rationality of the developed model and can provide valuable insights of safety regulation strategies for the government from the macro-level perspective.

Keywords: construction industry, accident situation, structure entropy weight method (SEWM), dynamic classification method (DCM), macro-level perspective.

Introduction

The construction industry always plays a significant role in promoting the development of national economy (Dong et al., 2015; N. Prasevic & Z. Prasevic, 2017; Zhou, Goh, & Li, 2015), but it, as one of the most hazardous sections, is responsible for 30–40% of fatalities in the world (Poh, Ubeynarayana, & Goh, 2018; Sunindijo & Zou, 2012; W. Zhang, X. Zhang, Luo, & Zhao, 2019). As shown on the web of the Ministry of Housing and Urban-Rural Development [MHURD] of China, the number of fatal accidents has a dramatic increase of 66.1% in the past few years, namely from 442 in 2015 to 734 in 2018 (MHURD, 2019a). Besides lots of casualties, enormous socio-economic consequences are always caused by construction accidents (Forteza, Carretero-Gomez, & Sese, 2017; Wang et al., 2018a). Generally, the accident occurrence is highly connected with the construction enterprise' effort in safety management and the government's effort in safety regulation (Hausken & Zhuang, 2016; Wang, Wu, Kang, Huang,

& Pan, 2018b; Wang et al., 2018d), while national/provincial safety regulation strategies influence the action of the enterprise in safety management and afterwards influence the accident rate in the construction industry (Ma & Zhao, 2018; Wang, Wu, Kang, Reniers, & Huang, 2018c). Poor accident situation of the entire industry has brought great safety regulation pressure to the government (e.g., national/provincial administrative department headers and ministerial safety department management personnel). Under such severe circumstances, it is urgent for the government to effectively lower accident/incident rates and improve construction accident situation.

Currently, many studies focused on the occurrence, evolution, assessment and prevention of accidents from the construction-site/project/accident level, such as construction safety management process (Wu, Liu, Zhang, Skibniewski, & Wang, 2015; Xiahou, Yuan, Li, & Skibniewski, 2018), the impact of individual and group char-

*Corresponding author. E-mail: zhigenhu@whu.edu.cn

acteristics on construction safety (Hasanzadeh, Esmaeili, & Dodd, 2017; E. Kim, Yu, K. Kim, & K. Kim, 2011), and construction accident causation mechanism (Fang, Ding, Luo, & Love, 2018; Zheng, Zhou, Wang, & Chen, 2018). Notably, previous research findings mostly benefit from detailed accident cause and scenario analysis, and can help safety management personnel of the enterprise (e.g., frontline supervisors, frontline safety managers and upper management personnel) understand and improve the accident situation at the construction-site level through ascertaining recommended measures with the potential to prevent construction accidents. But for the government, they can hardly obtain adequate information from those findings to grasp the accident situation of the entire construction industry, such as accident situations in different regions. Generally, the formulation of safety regulation strategies and the improvement of the regional/provincial accident situations benefit more from understanding macro-level accident situations in the building industry (Tam, Zeng, & Deng, 2004; Wang et al., 2018a; Yassin & Martonik, 2004). Thus, the government could pay more attention to accident situations of the building industry from the macro-level perspective. For instance, several Chinese ministerial administrative departments have set June as national safety production month since 2002, aiming at accident prevention and socio-economic sustainable development. During the period, safety-related activities should be more strictly conducted and more closely supervised from the national level, as well as stiffer penalties if accidents occurring. The fact that June hardly saw too poor accident situations is a proof of macro-level focus on accident situations. However, there are few studies on macro-level accident situations. This study aims to assess accident situations in China building industry from the macro-level perspective and provide valuable directions of safety regulation strategies for the government.

The accident situation is regarded as the barometer that can reflect the danger level in the construction industry (Coates, 2011). It is usually measured by some indicators related to construction accident rate such as the fatality rate per 100,000,000 yuan of gross domestic product (FGDP) (Shao, Hu, Liu, Chen, & He, 2019; Wang et al., 2018c) and the fatality rate per 100,000 construction practitioners (FCP) (Eteifa & El-adaway, 2018). Generally, these indicators can be used independently to assess/compare accident situations of the building industry, but the results vary with different indicators (Shao et al., 2019). For example, Beijing and Hubei experienced the FGDP of 1.366 and 0.841 in 2015 respectively, while the FCP of 0.013 and 0.035 for them respectively. These indicate that Beijing had worse accident situations than Hubei for FGDP, but Hubei had worse accident situations than Beijing for FCP. One-indicator-based results could mislead the government's decision-making in safety regulation. Furthermore, multi-dimensional characteristics of accident situations could be overlooked when only using one of these indicators. How to understand accident situations from multiple perspectives is very essential for the for-

mulation of safety regulation strategies, but the researches based on this idea are currently rare. Therefore, this study attempts to make it possible by establishing a comprehensive assessment model including various indicators related to accident situations.

Various evaluation methods can be implemented in the comprehensive assessment model of accident situations (Davoudabadi, Mousavi, Saparauskas, & Gitinavard, 2019). These methods can be divided into two kinds of qualitative and quantitative evaluation methods (Li et al., 2011), in which the sorting method is usually adopted by many researchers (Bao, Ruan, Shen, Hermans, & Janssens, 2012). This method can intuitively assess accident situations in different regions, but it cannot figure out characteristics of accident situations in different regions. Especially when involving too many regions, assessment results based on the method are deficient for the government to further explore the characteristics of different regions. Compared to the sorting method, the classifying method is another widely-used assessment technique that seeks to map accident elements into homogeneous clusters (Depaire, Wets, & Vanhoof, 2008; Hola & Nowobilski, 2018). Through the clustering analysis, some significant features of original data can be identified and mined in depth (Fahad et al., 2014). Generally, clustering algorithms can be classified into partitional methods and hierarchical methods (Horta & Camanho, 2014). Among them, automatically determining the number of clusters has been one of the most challenging problems in data classification (Jain, 2008). To solve the issue, this study proposed DCM based on clustering idea. This method is used to mine and clarify the characteristics of accident situations in the building industry, and it can reasonably determine the number of classifications by using an error function (S. Chen, Shao, Y. Chen, & Zheng, 2015). Notably, every indicator is considered equal weight in DCM, ignoring indicator weighting (Lin, 1989). However, the indicators related to accident situations contain different accident information. Generally, these indicators have different weights when conducting a comprehensive assessment on the accident situations. To address this issue, indicator weighting should be considered in DCM.

Currently, there are two main approaches of indicator weighting, namely subjective weighting and objective weighting (F. Liu, Zhao, Weng, & Y. Q. Liu, 2017). The former depends on the experts' experience and preference, such as *Delphi* survey and analytic hierarchy process (AHP) (Lai et al., 2015; N. Prasevic & Z. Prasevic, 2017; Zhao, Guo, Huang, & Zhong, 2017); the latter is applied in weight calculation by means of the original and objective data, such as principal component analysis and entropy weight (Cai et al., 2016; Wu, Wang, Z. P. Yang, Li, & Y. P. Yang, 2018). Yet both the two weighting techniques have their advantages and disadvantages. The subjective weighting owns highly explanation, but it could lack objectivity; the objective weighting can present higher accuracy, but it could be discordance with the actual conditions. To solve this problem, some researchers attempt to calculate indi-

cator weights through combining subjective weighting and objective weighting (Lai et al., 2015; Wu et al., 2018). Although this combination could, to some extent, integrate the advantages of the two traditional techniques, it may involve two aspects of shortcomings (G. Li, J. Li, Sun, & Zhao, 2017): (1) the combining formula lacks consistent and scientific expression (Dong, Shen, Jia, & Sun, 2018; Wu et al., 2018); (2) the calculation process of combined weight is too much complicated when conducting scientific deduction of the combining formula (Cui, Feng, Jin, & Liu, 2018; Dai & Niu, 2017; Lai et al., 2015). Considering that the rationality and simplicity of the combined weight cannot be addressed clearly, this study introduced SEWM to determine indicator weights. SEWM is developed based on the *Delphi* survey and entropy theory, and integrates subjective information and objective information in the process of weight calculation (Zheng, Shao, L. Chen, S. Chen, & Ge, 2014). This method cannot only simplify the calculation process, but also increase the reasonableness of weight determination (Liu et al., 2017; Zhao et al., 2017).

Therefore, this study is conducted to propose the comprehensive assessment model on accident situations in the building industry through establishing an assessment indicator system and integrating SEWM and DCM, aiming to provide targeted insights of safety regulation strategies for the government. The contribution of this study lies in three aspects: 1) establishing an assessment indicator system that can reflect accident situations of the building industry from different perspectives; 2) developing a comprehensive assessment model that can be used to classify the characteristic features of accident situations of the building industry; and 3) providing some insights for people who are concerned about safety regulation of the building industry from the macro-level perspective.

The rest of the paper is organized as follows. The research on indicators measuring accident situations is reviewed in Section 1. The comprehensive assessment model of accident situations is developed in Section 2, including the selection of indicators and the introduction of SEWM and DCM. Section 3 presents the study case and results. Result analysis and relevant discussion are conducted in Section 4, followed by Conclusions and future research.

1. Literature review

The accident situation is the barometer to reflect the extent of the danger related to the construction industry (Coates, 2011). A good understanding of macro-level accident situation can help explore potential practices of construction accident prevention (Yassin & Martonik, 2004). To figure out accident situations, the first step is to ascertain how to quantify them. Traditionally, the accident situation can be mostly measured by a series of basic indicators such as the number of accidents and the number of casualties (HOLA, 2009). However, it may vary with different non-accident indicators such as the economic development level and construction practitioners (Shao et al., 2019). For example, the US construction industry saw the highest number of

fatalities among industries, but the fourth highest fatality rate when considering the full-time equivalent worker (Karakhan, Rajendran, Gambatese, & Nnaji, 2018). Therefore, the basic indicators cannot be employed to sufficiently compare/assess relative level of accident situations.

To address the issue, some studies focus increasing on developing composite accident situation indicator (CASI), which is considered to be an analytical measurement approach for interpreting accident data (Coates, 2011). CASI is mostly a mathematical aggregation of two or more basic indicators that can reflect production situation of the construction industry. On the whole, CASIs are divided into three categories: accident frequency, accident severity and accident trend (Dong et al., 2011; Sari, Selcuk, Karpuz, & Duzgun, 2009; Soltanzadeh, Mohammadfam, Moghimbeygi, & Ghiasvand, 2017). From the accident frequency level, fatal injury rate per 100,000 full-time equivalent workers are widely used to estimate annual average fatalities based on construction time perspective (Marsh & Fosbroke, 2015; Mendeloff & Burns, 2013); fatal injury rate per 100,000 construction practitioners can be considered to explain the average fatalities of construction practitioners based on construction personnel perspective (Coates, 2011; Eteifa & El-adaway, 2018); fatal injury rate per 100,000,000 yuan of GDP in the construction industry is always selected to reflect the harmonious level between the construction industry and economic development based on construction output value perspective. From the accident severity level, Bellamy (2015) used the ratio of serious injuries to fatalities (RSF) to calculate the accident lethality, and the lower the RSF is, the more severity; Wang et al. (2018c) described fatality rate per one accident by the ratio of the fatality toll to the number of accidents (RFA), but the lower the RFA is, the less severity; Soltanzadeh et al. (2017) declared that accident severity rate can be measured by dividing working days-lost (multiplied by 2000) by all hours-worked considered. From the accident trend level, the ratio of the number of fatalities during one given period to that during another can be regarded as an expression of change trend of the number of fatal injuries, with the calculation similar to change trend of the number of accidents (Dong et al., 2011; HOLA, 2009; Kang, Siddiqui, Suk, Chi, & Kim, 2017).

Notably, the connotation of these CASIs are highly self-explanatory, and they can be independently used to assess accident situations of regions from the macro-level perspective (Wang et al., 2018c). However, assessment results of accident situations may be different when using different CASIs. Ambiguous or unreliable results could provide inappropriate decision-making information for the government (Wang et al., 2018c). In addition, only using one of these CASIs to evaluate accident situations may ignore the diversity of accident situation characteristics in one region. To solve the problem, a set of CASIs in the building industry can be considered for the comprehensive assessment of accident situations from the macro-level perspective. Based on China's national conditions, this paper aims to propose a comprehensive evaluation model

on the accident situation of the building industry through selecting the CASIs from different perspectives and help improve the understanding of China accident situations in construction safety regulation strategies.

2. Methodology and research process

This study firstly selected six CASIs from different perspectives to establish the assessment indicator system of accident situations based on China national conditions; afterwards, SEWM was applied to determine the weights of all indicators through comparing their impacts on accident situations; DCM was then proposed to mine and clarify the characteristics of accident situations in China building industry by means of classifying China provinces. The proposed comprehensive assessment model integrating SEWM and DCM is depicted in Figure 1 and in following subsection.

2.1. Establishing assessment indicator system

To explore the characteristics of accident situations in China provinces, the first step is to select reasonable indicators that can be used to measure accident situations in China building industry. Following a wide review of the literature and interviews with safety experts, various indicators related to construction accident rate have been used to assess accident situations of the building industry in many countries. However, each country generally chooses different indicators based on its own national condition. For instance, the fatality rate per 100,000 full-time equivalent workers is widely adopted in many developed countries, but it is not a statistical indicator in China due to its current condition in the building industry (Shao et al., 2019). Considering the availability and acceptable quality of original data of the indicators at the current development stage of China, six indicators measuring accident situations were chosen from three perspectives includ-

ing fatal accident frequency (FAF), fatal accident severity (FAS) and fatal accident trend (FAT), as listed in Table 1.

Among the six indicators, there are three indicators that reflect the fatal accident frequency in the building industry, namely the fatality rate per 100,000 practitioners (IS1), the fatality rate per 1,000,000 m² of floor areas (IS2), fatality rate per 100,000,000 yuan of GDP (IS3). The IS1 is widely applied in many countries (Irumba, 2014) and indicates the level of occupational risks associated with the implementation of construction works (Hola & Szostak, 2015). Due to the rapidly growing development of China real estate industry in recent years, the IS2 is regarded as a relatively scientific indicator and has been applied in China building industry (Ma, Chen, & Liu, 2015). The IS3 is an important indicator to measure whether a region actually integrates safety production policies into the overall planning of regional economic and social progress and strictly implements them (Liu & Wu, 2011). It indicates the level of production safety under certain economic conditions in the building industry. Moreover, there is one indicator that reflects the fatal accident severity in the building industry, namely fatality rate per one fatal accident (IS4). The IS4 can present the average lethality of fatal accidents (Wang et al., 2018c). At last, the rest indicators can reflect the fatal accident trend in the building industry, namely the change trend of the number of fatal accidents (IS5) and the change trend of the number of fatalities (IS6). They are frequently used to describe temporal characteristics of accident situations in given regions (Marsh & Fosbroke, 2015). The calculation method of selected indicators is shown in the fourth column of Table 1.

1.1. Weighting indicators based on structure entropy weight method

Determining the weights of indicators is the key to the comprehensive assessment of accident situations. The reasonableness of weight determination directly affects the

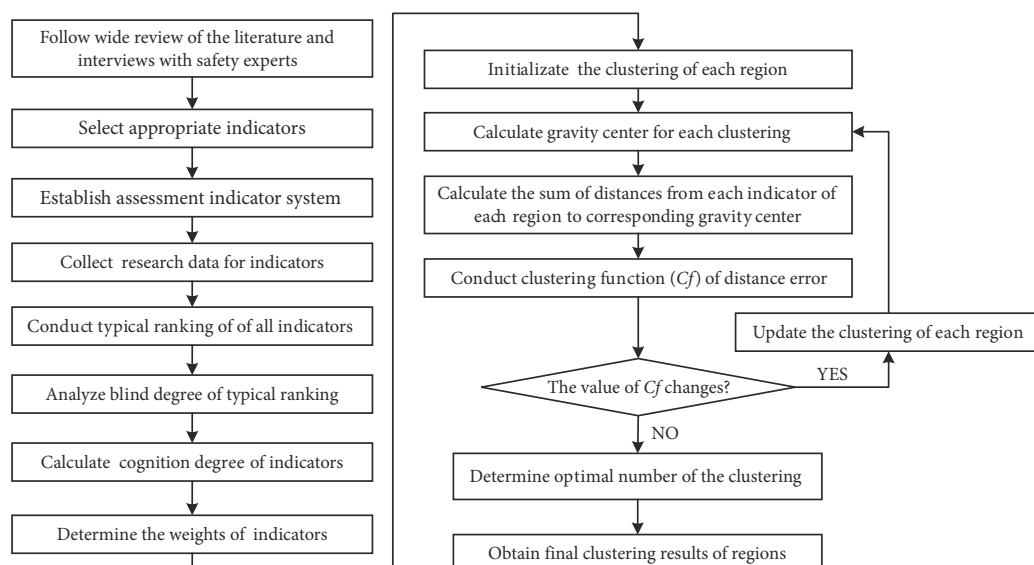


Figure 1. Framework flowchart of comprehensive assessment model on accident situations of the construction industry

Table 1. Indicators measuring accident situations

| Perspective | Indicator | Symbol | Calculation method |
|-------------|---|--------|---|
| FAF | Fatality rate per 100,000 practitioners | IS1 | Fatalities / Construction practitioners (one hundred thousand persons) |
| | Fatality rate per 1,000,000 m ² of floor areas | IS2 | Fatalities / Floor areas (one million square meters) |
| | Fatality rate per 100,000,000 yuan of GDP | IS3 | Fatalities / GDP of building industry (one hundred million yuan) |
| FAS | Fatality rate per one fatal accident | IS4 | Fatalities / Fatal accidents |
| FAT | Change trend of the number of fatal accidents | IS5 | (Fatal accidents (this year) – Fatal accidents (last year)) / Fatal accidents (last year) |
| | Change trend of the number of fatalities | IS6 | (Fatalities (this year) – Fatalities (last year)) / Fatalities (last year) |

Note: The indicators are selected from representative literature (Liu & Wu, 2011; Hola & Szostak, 2015; Irumba, 2014; Ma et al., 2015; Marsh & Fosbroke, 2015; Wang et al., 2018c).

merits of assessment results. SEWM is a method combining qualitative analysis and quantitative analysis. Its basic idea is to gather experts' typical ranking related to the indicators, and calculate their entropy value based on entropy theory. Meanwhile, this method can eliminate the uncertainty in the indicator cognition by the blind degree analysis, and retain the reasonable part of the subjective influence (Ding, Chen, Cheng, & Wang, 2015; Liu et al., 2017). The calculation steps are as follows:

Step 1: Collecting experts' opinions for typical ranking

Based on *Delphi* survey, a questionnaire form is designed for experts to determine qualitative rankings of indicators. Notably, the experts are invited to anonymously fill in the questionnaires in groups, with ranking the relative importance of indicators by means of to their knowledge and experience. More important the indicator is, more forward its ranking is. Besides, some indicators could be recognized as equally important and they would have the same ranking. Considering the differences of experts' cognitions on the indicators, several experts from different knowledge background are generally assembled a group to decide the final ranking of the indicators, which is called "typical ranking".

Step 2: Analyzing the blind degree of typical ranking

To reduce the uncertainty of expert ranking (blind degree), the entropy value was calculated based on the entropy theory. Assuming there are p expert groups participating in the survey of n indicators, which is marked as $Y = \{Y_1, Y_2, \dots, Y_k, \dots, Y_n\}$, ($k \in (1, 2, \dots, n)$), all the questionnaires can be recovered. The typical ranking from sth expert group is recorded as $\{I_{s1}, I_{s2}, \dots, I_{sk}, \dots, I_{sn}\}$, ($s \in (1, 2, \dots, p)$), and I_{sk} stands for the ranking number of experts' judgment on indicator Y_k , which can be any natural number that is no more than n . For p expert groups, the matrix of typical rankings can be expressed as follows:

$$H = (I_{sk})_{p \times n} = \begin{bmatrix} I_{11} & I_{12} & \cdots & I_{1n} \\ I_{21} & I_{22} & \cdots & I_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ I_{s1} & I_{s2} & \cdots & I_{pn} \end{bmatrix}. \quad (1)$$

Then the typical rankings can be transformed into quantitative entropy values through the membership function $e(I)$:

$$e(I) = -\lambda u_k(I) \ln u_k(I). \quad (2)$$

Setting $u_k(I) = (\theta - I) / (\theta - 1)$ and $\lambda = 1 / \ln(\theta - 1)$, the Eqn can be derived as:

$$e(I) = \frac{\theta - I}{\theta - 1} \left(1 - \frac{\ln(\theta - I)}{\ln(\theta - 1)} \right). \quad (3)$$

Assuming $1 - e(I) / \left(\frac{\theta - I}{\theta - 1} \right) = g(I)$, then

$$g(I) = \ln(\theta - I) / \ln(\theta - 1). \quad (4)$$

In above formulas, θ stands for the conversion parameter, which is defined as $\theta = n + 2$. $g(I)$ is the membership function of ranking number I . When putting the qualitative ranking number I_{sk} into the formula, the quantitative converted value of I_{sk} can be obtained by $b_{sk} = g(I_{sk})$, which is the membership degree of ranking number I_{sk} . Correspondingly, the matrix $B = (b_{sk})_{p \times n}$ is defined as the membership degree matrix:

$$B = (b_{sk})_{p \times n} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{s1} & b_{s2} & \cdots & b_{pn} \end{bmatrix}. \quad (5)$$

Then the average cognition degree on the indicator Y_k from p expert groups, marked as b_k , is calculated by the Eqn (6), which can reflect experts' consistent opinions:

$$b_k = (b_{1k} + b_{2k} + \cdots + b_{pk}) / p. \quad (6)$$

The uncertainty caused by experts' cognition on Y_k is defined as cognition blind degree Q_k . Obviously, $Q_k > 0$.

$$Q_k = \left[\max(b_{1k}, b_{2k}, \dots, b_{pk}) + \min(b_{1k}, b_{2k}, \dots, b_{pk}) \right] / 2 - b_k. \quad (7)$$

Step 3: Calculating the overall cognition degree of all indicators

For Y_k , the overall cognition degree v_k can be calculated by the Eqn (8). Therefore, $\mathbf{v} = (v_1, v_2, \dots, v_k, \dots, v_n)$ represents overall cognition degrees for all indicators, respectively:

$$v_k = b_k(1 - Q_k). \quad (8)$$

Step 4: Determining the weights of all indicators

Normalizing $\mathbf{v} = (v_1, v_2, \dots, v_k, \dots, v_n)$ by means of the Eqn (9), the weight of each indicator can be obtained, namely the weight vector $\mathbf{w} = (w_1, w_2, \dots, w_k, \dots, w_n)$ and

$$\sum_{k=1}^n w_k = 1. \quad (9)$$

$$w_k = v_k / \sum_{k=1}^n v_k.$$

2.3. Clustering provinces based on dynamic classification method

The proposed DCM is a quantitative multi-indicator numerical classification method, which is applied in the classification of rock stability for the first time (Lin, 1989). The basic idea is “cluster analysis” that considers the sample as a point in classification space. According to the specific regulation, the nearest points of the “distance” are classified through repeated iterations, and the classification results of the sample are obtained when the error is the smallest. The reason why DCM is used to classify the accident situations in China building industry is: First, previous researches generally depend on the known grading standards, and DCM can cluster the data in the case of unknown grading standards. Second, the method can make full use of the objective historical data of building construction accidents, reflecting the objectivity of grading. The detailed steps are as follows.

Step 1: Distance definition

The calculated parameters form the basis for classifying provinces with the use of cluster analysis. To perform the classification, it is necessary to measure the distances between the assessment samples. This study uses Euclidean distance d_{ij} to characterize the difference between the i th sample and the j th sample:

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2}, \quad (10)$$

where n represents the number of indicators, x_{ik} represents the characteristic value of the indicator Y_k of the i th sample.

Step 2: Normalized treatment

Due to the different measured scales of the selected indicators, the data need be normalized to assure good comparability between indicators. Meanwhile, to reflect the different influence of different indicators on the assessment objects, it is necessary to rationally allocate indica-

tor weights. Thus, the raw data is processed by the range method, combined with the weight of each indicator. The characteristic values of the indicators of all samples are normalized as follows:

$$x'_{ik} = w_k \frac{x_{ik} - \min \mathbf{x}_k}{\max \mathbf{x}_k - \min \mathbf{x}_k}, \quad (11)$$

where x'_{ik} is the characteristic value of Y_k for the i th sample after normalized treatment, x_{ik} is the original value of Y_k for the i th sample, \mathbf{x}_k is the vector of the original value of Y_k for all samples.

Step 3: Clustering iteration

1) The number of clusters. As mentioned above, DCM is suitable for the situation where there is no clear grading standard, so the determination of the optimal classification number needs to be analyzed by comparing different classification results. The number of clusters is often determined according to classification requirements of the actual situation, and it should not be too detailed or too rough, because it is too detailed to manage and too rough to have practical significance.

2) Clustering initialization. Firstly, calculate the composite value of the i th sample:

$$s(i) = \sum_{k=1}^n x'_{ik}. \quad (12)$$

Then, when the number of clusters is D , the initial clustering of the i th sample belongs to:

$$N(i) = \text{Fix} \left\{ \frac{(D-1) \left[\max_{1 \leq i \leq m} s(i) - s(i) \right]}{\max_{1 \leq i \leq m} s(i) - \min_{1 \leq i \leq m} s(i)} + 0.5 \right\} + 1, \quad (13)$$

where $\text{Fix}()$ represents the rounding function.

3) Calculation of gravity center for each clustering. The gravity center of the k th indicator of the r th clustering is calculated as follows:

$$C_{r,k} = \sum_{i=1}^n \delta x'_{ik} / \sum_{i=1}^n \delta, \quad \delta = \begin{cases} 1, N(i) = r \\ 0, N(i) \neq r \end{cases}. \quad (14)$$

4) Clustering update. Calculate the sum of distances from indicators of the i th sample to $C_{r,k}$:

$$d_{i,r} = \sqrt{\sum_{k=1}^n (x'_{ik} - C_{r,k})^2}. \quad (15)$$

Update the clustering of the i th sample according to the proximity principle:

$$N'(i) = l, l \in \{1, 2, \dots, r\}; d_{i,l} = \min(d_{i,1}, d_{i,2}, \dots, d_{i,r}), \quad (16)$$

where $N'(i)$ is the new clustering of the i th sample.

5) Iteration output. Repeat the processes of 2–4 and constantly compare the difference value of the gravity center before and after the iterations. If the value is smaller than the specified error value or the value is no longer changing, the final clustering result will be output.

Step 4: Determination of optimal number of the clustering

The sum of the distances from all samples to the gravity center of their clustering is defined as a clustering function C_f , which is essentially an error function:

$$C_f = \sum_{i=1}^n d_{i,r}. \quad (17)$$

The optimal number of the clustering can be determined by comparing the change of the function value with respect to the number of clusters.

3. Comprehensive assessment on accident situations of China provinces

According to the MHURD of China, the housing and municipal construction industry encounters continuous increase in construction accidents, and 2018 experienced the highest number of fatalities (840) in recent five years, with daily 2.3 fatalities. Poor accident situation has posed a great challenge to the government's safety supervision. Effective safety regulation strategies are thus much urgent. In the case study of China provinces, this section presents the application of the comprehensive assessment model on accident situations, aiming at providing valuable insights of safety regulation strategies from the macro-level perspective.

According to the calculation method of each indicator in Table 1, the data to be collected includes the number of fatal accidents, the number of fatalities, the building construction population, the floor areas and the GDP of the building industry in each province. Considering the availability and acceptable quality of original data, the data in 2015 was applied in the comprehensive assessment model on accident situations in China building industry. Among them, accident data of each province in 2015 (including the number of fatal accidents and the number of fatalities) came from MHURD (2019b), which conducts short reports of construction fatal accidents in China. The rest data came from the National Bureau of Statistics [NBS] of China (NBS, 2019), which is mainly responsible for publishing statistical data related to China social and economic development. Considering the deficiency or incompleteness of the data from Tibet, Hong Kong, Macao and Taiwan, the data in 30 Chinese provinces were finally used for detailed analysis in this study.

3.1. Determination of indicator weights

It should be noted that the selection of safety experts is the key to reasonably determine the weight of indicators. In this study, the selected experts should work for over five years and perform well in the field of construction safety management/research. When conducting the importance ranking of indicators that can measure accident situations in the building industry, totally 12 experts were invited to anonymously fill in the typical ranking on all indicators in Table 1. These experts came from Wuhan University (2),

China Three Gorges University (2), Hubei Anyuan Safety and Environmental Protection Technology Company Limited (3), Wuhan Hanyang Municipal Construction Group Corporation (3), Changjiang Institute of Survey, Planning, Design and Research (2), respectively. To reasonably and efficiently offer quantitative judgments on the importance of the indicator set, these experts were divided into four groups, with three experts in each group. Based on the consultation and discussion of each group, the ranking results from the four expert groups were collected in the second-fifth columns of Table 2. Eventually, the weights of the six indicators were obtained according to the calculation steps, as listed in the last column of Table 2.

Table 2. Calculation of the weights of indicators

| Indicator | Typical ranking | | | | Weight |
|-----------|-----------------|---------|---------|---------|--------|
| | Group 1 | Group 2 | Group 3 | Group 4 | |
| IS1 | 2 | 2 | 2 | 3 | 0.1879 |
| IS2 | 6 | 5 | 4 | 4 | 0.1192 |
| IS3 | 4 | 4 | 4 | 5 | 0.1395 |
| IS4 | 1 | 1 | 1 | 2 | 0.2060 |
| IS5 | 5 | 3 | 5 | 1 | 0.1516 |
| IS6 | 3 | 1 | 3 | 1 | 0.1959 |

3.2. Clustering results of provinces

To explore the characteristics of the accident situations in Chinese provinces, DCM was used to classify provinces with regards to their accident situations. The first step of the classification was to determine the number of clusters, but the number should not be too small or big when involving a large database (Raviv, Fishbain, & Shapira, 2017). Therefore, this study firstly assumed that the clustering number D ranges from 2 to 9 (C2–C9). According to the calculation process in Subsection 2.3, the clustering results in different values of D were obtained by using the Matlab software, as shown in Figure 2. It can be seen that the classification results varied with different values of D , but their change trends were highly consistent.

Meanwhile, the values of the clustering function corresponding to different classification number were calculated, as presented in Figure 3. The figure showed that the C_f function (error function) was well fitted by the function $y = 0.2223x^{-0.6557}$, and the power function indicated that the C_f values (classification errors) decreased moderately as the classification number increased, namely smaller and smaller classification errors. Especially when the classification number was big enough, the classification error could gradually approach zero, which accords with the actual situation.

Moreover, the reduction percentages of the classification errors were considered to determine the appropriate number of clusters. As shown in Figure 3, the curve of the reduction percentage had the smallest value when the classification number was only equal to four. Meanwhile, the absolute value of the slope of the fitting curve when

$2 \leq D \leq 4$ was too much bigger than that when $4 \leq D \leq 9$, which meant that the clustering effect would not change much after $D = 4$. Based on above facts, four clusters were the most appropriate to minimize the outliers and classify each cluster efficiently. Eventually, the classification results were presented in Table 3.

4. Discussion

The result reliability is the key to the comprehensive assessment model of accident situations. Indicator weights for SEWM and clustering results for DCM are thus validated by the actual condition of the building industry.

Result analysis is further carried out to clarify the characteristics of accident situations in China building industry. The findings can be significant references to safety regulation strategy for the government.

4.1. Method validation

4.1.1. Indicator weighting

There are many methods that can calculate indicator weights. In this study, SEWM is used to determine the weights of indicators in China provinces. Considering that the data in this study can be used to calculate entropy weights of indicators, only the entropy weight method

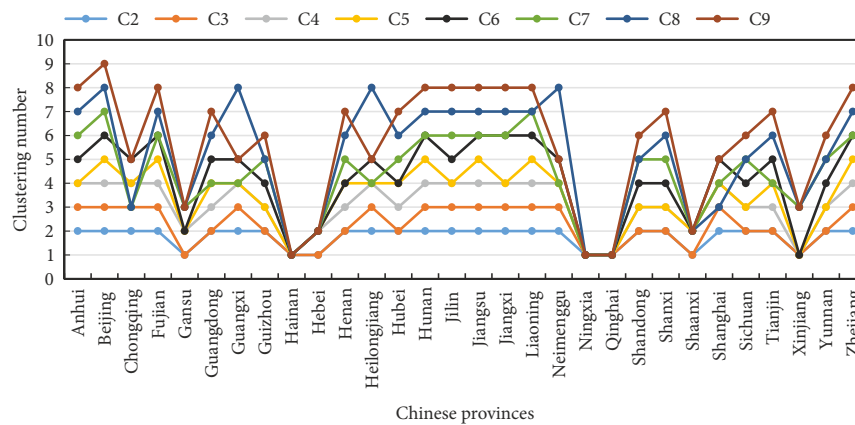


Figure 2. Dynamic classification results corresponding to different classification number

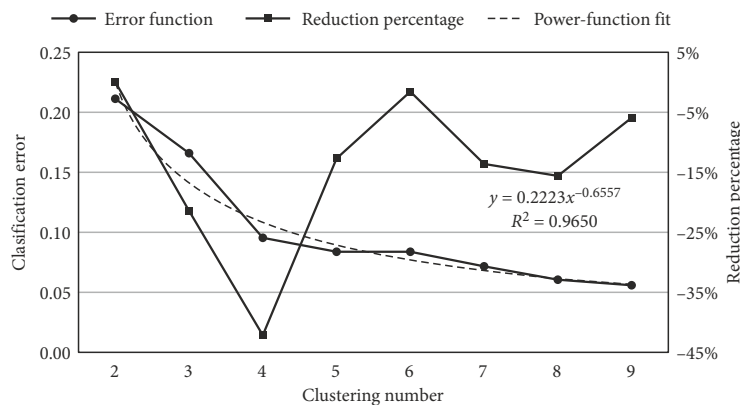


Figure 3. Error function diagram

Table 3. Classification results of provinces for $D = 4$

| Province | Classification | Province | Classification | Province | Classification |
|-----------|----------------|-----------|----------------|--------------|----------------|
| Hainan | Cluster 1 | Hubei | Cluster 3 | Guangxi | Cluster 4 |
| Ningxia | Cluster 1 | Shandong | Cluster 3 | Heilongjiang | Cluster 4 |
| Qinghai | Cluster 1 | Shanxi | Cluster 3 | Hunan | Cluster 4 |
| Xinjiang | Cluster 1 | Sichuan | Cluster 3 | Jiangsu | Cluster 4 |
| Gansu | Cluster 2 | Tianjin | Cluster 3 | Jiangxi | Cluster 4 |
| Hebei | Cluster 2 | Yunnan | Cluster 3 | Jilin | Cluster 4 |
| Shaanxi | Cluster 2 | Anhui | Cluster 4 | Liaoning | Cluster 4 |
| Guangdong | Cluster 3 | Beijing | Cluster 4 | Neimenggu | Cluster 4 |
| Guizhou | Cluster 3 | Chongqing | Cluster 4 | Shanghai | Cluster 4 |
| Henan | Cluster 3 | Fujian | Cluster 4 | Zhejiang | Cluster 4 |

(one of the most common objective weighting) is chosen to compare with SEWM. Therefore, the objective weights of indicators are obtained by the entropy weight method according to the original data. The weighting results in the two methods, namely SEWM and entropy weight, are listed in Table 4.

It can be observed that weighting results are quite different. In reality, the fatality rate per one fatal accident (IS4) is an indicator that people can intuitively feel by describing the fatal accident severity, which is the most direct appearance of the accident situation in the building industry. Generally, the IS4 can be greater significance on the comprehensive assessment results of accident situations compared to other indicators. Moreover, the fatality rate per 1,000,000 m² of floor areas (IS2) is, to some extent, a controversial indicator in China, which indicates that the IS2 should not have a dominant influence on assessment results (Yuan, 2005). Only based on the two indicators, SEWM weights are highly in line with the actual situation, but the entropy weights not. Therefore, the entropy weight method is not suitable for this study, although it makes indicator weights more objective. Compared to the entropy weights, SEWM weights of the two indicators are more reasonable.

Besides, the range from 0.1192 to 0.2060 suggests that the difference of the impacts of SEWM weights on accident situations is not too big compared with entropy weights. It is a rather proof that the six indicators are always selected to assess the accident situations in the building industry despite different assessment results. AHP is one of the most common subjective weighting, but the consistency test can hardly be satisfied when involving too many indicator elements (Mahmoudzadeh & Bafandeh, 2013). Due to no use of the consistency test, SEWM has more concise computing process compared to AHP. Therefore, SEWM can be considered more suitable for indicator weighting in this study.

Table 4. Comparison of weighting results between SEWM and entropy weight.

| Weighting method | IS1 | IS2 | IS3 | IS4 | IS5 | IS6 |
|------------------|--------|--------|--------|--------|--------|--------|
| SEWM | 0.1879 | 0.1192 | 0.1395 | 0.2060 | 0.1516 | 0.1959 |
| Entropy weight | 0.1871 | 0.2681 | 0.2481 | 0.1405 | 0.0726 | 0.0836 |

Table 5. Comparison of classification results between DCM and K-means

| Classification | Clustering method | |
|----------------|--|---|
| | DCM | K-means method |
| Cluster 1 | Hainan, Ningxia, Qinghai, Xinjiang | Hainan, Ningxia, Qinghai, Xinjiang |
| Cluster 2 | Gansu, Hebei, Shaanxi | Gansu, Hebei, Shaanxi |
| Cluster 3 | Guangdong [#] , Guizhou, Henan [#] , Hubei [#] , Shandong, Shanxi [#] , Sichuan, Tianjin [#] , Yunnan | Guizhou, Shandong, Sichuan, Yunnan |
| Cluster 4 | Anhui, Beijing, Chongqing, Fujian, Guangxi, Heilongjiang, Hunan, Jilin, Jiangsu, Jiangxi, Liaoning, Neimenggu, Shanghai, Zhejiang | Anhui, Beijing, Chongqing, Fujian, Guangdong [#] , Guangxi, Heilongjiang, Henan [#] , Hubei [#] , Hunan, Jilin, Jiangsu, Jiangxi, Liaoning, Neimenggu, Shanghai, Shanxi [#] , Tianjin [#] , Zhejiang |

Note: [#]indicates that the provinces have different classification in two methods.

4.1.2. Classification of provinces

In this study, DCM is proposed to classify Chinese provinces that have different accident situations in the building industry. Notably, classification results by using various methods could never be identical (Wegman & Oppe, 2010). The K-means method, one of the most commonly clustering methods, is adopted to verify the rationality and applicability of DCM, which is a quite necessary job. As shown in Table 5, the classification results from DCM and K-means method are highly similar. On the whole, the similarity of their results reaches 83.3%. Especially for Cluster 1 and Cluster 2, the results of DCM are the same as that of K-means method. The only difference is the case that five provinces including Guangdong, Henan, Hubei, Shanxi and Tianjin, belong to Cluster 3 for DCM, but Cluster 4 for K-means method.

The characteristics of accident situations that can be reflected by the values of the gravity center for different clusters are obtained, as shown in Figure 4. Obviously, Cluster 1 experienced higher FAF, while higher FAT for Cluster 2

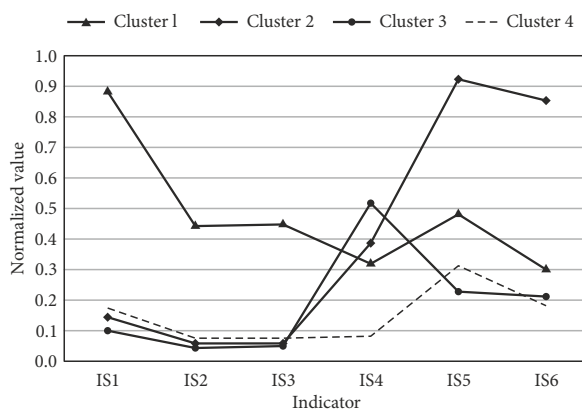


Figure 4. The characteristics of accident situations in different clusters

and higher FAS for Cluster 3. On the whole, Cluster 4 reflects relatively good accident situations compared to other clusters. As for the above-mentioned five provinces, they highly accord with the characteristics of Cluster 3. There are two levels of reasons: 1) for the five provinces, their IS4 (1.5000, 1.5000, 1.4667, 1.3889 and 1.3750) are more than the average IS4 in all provinces (1.3408), which indicates that the five provinces have higher FAS than entire average level; 2) for Cluster 3, when the five provinces are not considered, the average IS4 is 1.8806 for Cluster 3 and 1.1019 for Cluster 4, which implies that Cluster 3 is characterized by higher FAS and Cluster 4 by lower FAS compared to entire average level. It is thus objective that the five provinces should belong to Cluster 3 rather than Cluster 4. Compared to K-means method, clustering results of DCM are, to a large extent, more suitable for the actual condition. Therefore, proposed DCM can be regarded as a reasonable classification method.

4.2. Result analysis

This study classified Chinese provinces through establishing the assessment indicator system and integrating SEWM and DCM. The results show that 30 provinces in China are classified into four clusters based on their accident situations. Moreover, the characteristics of accident situations in different clusters are further explored to clarify the similarities and differences within and between clusters at a provincial level. The average values of indicators in different clusters are presented in Table 6.

Obviously, there is a big difference in the distribution of the averages in different clusters. Cluster 1 provinces see higher FAF including IS1, IS2, and IS3, which are over four times more than that of other clusters, respectively. In other words, poor accident situations in Cluster 1 provinces mainly result from the higher FAF. Notably, the three FAF indicators are the mathematical aggregation of basic indicators such as the number of fatalities and GDP, and their values were easily influenced by that of basic indicators, but not necessarily (Wang et al., 2018c). Taking IS3

for example, if the fatalities and GDP remained the same proportional change, IS3 would be unchanged. Moreover, practitioners, floor areas and the GDP of the building industry have no significant linear correlation with IS1, IS2, and IS3 respectively (-0.492 , -0.356 , -0.369), but they have the most backward rankings in Cluster 1 provinces except Xinjiang, as shown in Table 7. This implies that one province could encounter higher FAF when the three factors of the province rank the bottom in all provinces. However, to some extent, GDP can influence the level of IS3 for one region, but the influence mechanism is very complicated due to so many factors and their relationship involved (Liu & Wu, 2011). There exist similar situations for practitioners and IS1, and floor areas and IS2. Therefore, it's difficult for Cluster 1 provinces to give a plausible reason why they have higher FAF, maybe for the small economic volume or for others. But one thing for sure is that the government should balance the development of the economy and the construction industry, introduce advanced construction techniques and learn advanced safety management methods from better-performing regions.

Furthermore, Cluster 2 provinces have higher FAT including IS5 (86.67%) and IS6 (160.00%). The shocking increases should arouse the vigilance from other provinces, although the IS5 and IS6 have overall declining rates of 16.02% and 15.78% respectively. The growth percentages even indicate that accident situations in Cluster 2 provinces are almost out of control compared to that last year. Considering that the building industry is the pillar section of the economic development, some provinces may radically pursue rapid growth of GDP at the cost of construction safety. However, the above economic development idea could be inconsistent with the actual situation. Especially for Hebei, its IS5 and IS6 have big growth, but its GDP has a declining rate of 7.48% compared to that last year. The same situation is, to a certain extent, suitable for Gansu ($-0.97%$) and Shaanxi ($-0.15%$). The facts could be caused by two reasons: one is that current construction safety regulations have serious drawbacks in these provinces;

Table 6. Distribution of average values of indicators in each cluster

| Classification | Number | IS1 | IS2 | IS3 | IS4 | IS5 | IS6 | Feature |
|----------------|--------|--------|--------|--------|--------|---------|---------|-----------|
| Cluster 1 | 4 | 5.9934 | 0.3454 | 0.0294 | 1.3983 | 12.59% | 7.33% | High FAF |
| Cluster 2 | 3 | 1.2400 | 0.0586 | 0.0052 | 1.4833 | 86.67% | 160.00% | High FAT |
| Cluster 3 | 9 | 0.9728 | 0.0472 | 0.0047 | 1.6392 | -30.94% | -17.16% | High FAS |
| Cluster 4 | 14 | 1.4350 | 0.0696 | 0.0062 | 1.1019 | -15.95% | -23.85% | Normality |
| Total/Average | 30 | 1.8846 | 0.0985 | 0.0087 | 1.3408 | -16.02% | -15.78% | - |

Table 7. Characteristics of Cluster 1 provinces

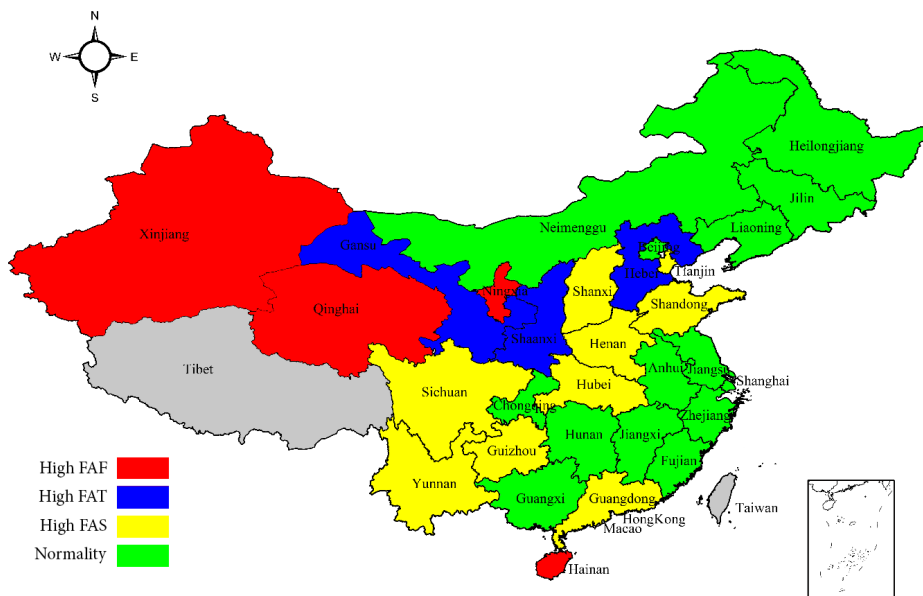
| Factor | Ranking | | | | Pearson correlation | | |
|--------------|---------|---------|---------|----------|---------------------|--------|--------|
| | Hainan | Ningxia | Qinghai | Xinjiang | IS1 | IS2 | IS3 |
| Practitioner | 30/30 | 28/30 | 29/30 | 26/30 | -0.492 | | |
| Floor area | 29/30 | 28/30 | 30/30 | 24/30 | | -0.356 | |
| GDP | 29/30 | 28/30 | 30/30 | 21/30 | | | -0.369 |

the other is that these provinces loosely implement existing construction safety policies. Therefore, the first step towards improving the poor accident situations in Cluster 2 provinces should be to focus on the supervision level of construction administrative departments and clarify the actual reason of percentage increases, avoiding the vicious circle of poor situations in the long term.

Moreover, Cluster 3 provinces have experienced higher FAS (IS4) of 1.64 deaths, which is approximately 50% more than that of Cluster 4 provinces. There are two main situations for the serious outcome: 1) personal death is immediately caused by construction accidents; 2) the injured persons whose deaths should have been avoided eventually die after the accidents due to inappropriate first-aid measures. For the first situation, those construction activities which could cause serious deaths (e.g., excavation of deep foundation pit, installation and demolition of large lifting machinery) should not be conducted until the allowance of expert demonstration. For example, the installation of the tower crane generally involves a certain number of persons who are working at height, and the activities probably pose safety hazards to these workers, even threatening other persons around the tower crane. The scheme of the installation thus has to be demonstrated in detail. Therefore, the government should think about how to strictly implement demonstration procedures of highly serious construction activities in Cluster 3 provinces. For the second situation, policies of accident emergency rescues should be further improved to avoid more dispensable casualties. Considering that the IS4 is an indicator that people can intuitively feel by using accident casualties in the building industry, the government should pay more attention to Cluster 3 provinces and reduce the severity of accidents, avoiding negative effects on industry reputation and social stability.

Additionally, the rest of provinces seem to have no salient features except the lowest IS4 and IS6, which represent relatively good/normal accident situations. Considering that Cluster 4 provinces account for about half of all provinces (14/30), the normal/good accident situations could be, to a certain extent, regarded as the most common in the Chinese provinces. Although Cluster 4 provinces could provide successful experience for other worse-performing provinces in China, they still have much more room for the improvement of accident situations compared to some developed countries.

According to different characteristics of accident situations in different clusters, the four clusters are defined as high FAF, high FAT, high FAS and normality respectively, as listed in the last column of Table 6. Additionally, accident situations in different clusters also present the characteristics of regional distribution at a provincial level, as shown in Figure 5. It is clearly found that provinces with normal/good accident situations are mainly distributed in the southeastern and northeastern regions of China. The worse-performing provinces with high FAF and high FAT are mainly distributed in the northwest region of China. Provinces with High FAS are mainly distributed in the southwestern and central regions of China. Therefore, the government should focus on accident situations of the building industry in the central and western regions of the underdeveloped economy. Especially at the national level, progressive policies such as resource allocation and economic development strategies should be inclined to these regions, with inspiring participants in the building industry to improve safety management level together. Furthermore, accident situations in the provinces vary with different safety management levels, and the better-performing provinces can generally provide successful experience for other worse-performing provinces in China.



Note: Tibet, Hong Kong, Macao and Taiwan were not considered in this study due to their incompleteness or/and deficiency of the data. Data source: National Bureau of Statistics (NBS, 2019), Ministry of Housing and Urban-Rural Development (MHURD, 2019b).

Figure 5. Regional distribution of accident situations in different clusters

Therefore, regional information sharing mechanism of safety regulation strategies, especially for what has proved to be successful in one province/country, should be advocated around the world, and more interaction and exchange of accident situations should be made between academia and industry.

Conclusions and future research

Due to the inherently hazardous nature, the construction industry experiences the bad record of the high accident rate in many countries, especially in China. Severe accident situation has posed serious challenges to the government safety regulation in the recent years. This study develops a comprehensive assessment model based on SEWM and DCM to clarify the characteristic features of accident situations in the building industry, aiming at providing valuable insights of safety regulation strategies from the macro-level perspective. The model is applied to assess accident situations of the China building industry. Through result analysis and discussion, the following findings can be concluded:

1. Based on China national condition, six indicators measuring accident situations are selected to establish an assessment indicator system. These indicators reflect the accident situations of the China construction industry from FAF, FAS and FAT perspectives, respectively. Among them, fatality rate per one accident (IS4) has the biggest representativeness on the accident situation, followed by the change trend of the number of fatalities (IS6).
2. SEWM and DCM used in this study are validated by the actual condition of China. Compared to traditional indicator weighting such as AHP and entropy weight method, SEWM is more suitable for this study due to concise calculation process and practical interpretation of weights. Meanwhile, clustering results of DCM are considered more reasonable compared with the K-means method. On the whole, the proposed model is feasible and reliable.
3. Accident situations in Chinese provinces present various features, namely high FAF, high FAT, high FAS and normality. Notably, the provinces with poor accident situations account for over half of all provinces (53.33%), and they are mainly distributed in the central and western regions of China where there exists the underdeveloped economy. Besides, provinces in the southeastern and northeastern regions of China perform better, but they still have much more room for the improvement of accident situations. Furthermore, the government should advocate regional information sharing of safety regulation at the national/provincial level.
4. The provinces that experience the lowest number of practitioners, floor areas and the GDP of the building industry could be prone to high FAF, especially for Hainan, Ningxia and Qinghai. The government should pay greater attention to these regions that have smaller scale of the construction industry in China and allocate

resources to balance the development of the economy and the construction industry.

5. Poor accident situations in the provinces with higher FAT, namely Gansu, Hebei and Shaanxi, could be almost out of control, with the fact that the number of fatal accidents still have a big growth despite the decline of GDP. As for these regions, the supervision level of construction administrative departments is, to some extent, more urgent to improve, and/or safety management strategies that could be currently missing must be supplemented in time through learning from better-performing regions.
6. FAT is a perspective that can show the most intuitively accident situation, thus the provinces with higher FAT should be paid more attention to safety warning and/or emergency plans before and after one accident occurring, such as irregularly publishing the lists and hazardous factors of those construction activities which could cause serious deaths.

This study may have some research limitations. Research results comply with current national conditions in China, but they are generally influenced by some factors such as the number of selected provinces/regions and the change of indicators. Thus the results could change if other regions (e.g., Hong Kong and Taiwan) or/and other indicators (e.g., the fatality rate per 100,000 full-time equivalent workers) are considered in the future work. Meanwhile, the comprehensive assessment model has some extensibility/flexibility, namely the indicators and regions contained can be increased or reduced based on the actual conditions of accident situations in the assessment regions/countries. Therefore, the model needs appropriate improvement when applied in other regions/countries. Moreover, the model is used to analyze static characteristics of provincial distribution of accident situations based on the one-year data (2015), but the accident situation is generally regarded as a dynamic phenomenon. Therefore, future work can be conducted to explore the dynamic property of the accident situation by using two- or more-year data. Furthermore, the influence mechanism of GDP on IS3 has been the research hotspot, but it is highly complicated because GDP is not the only factor affecting IS3. This kind of research topics can be conducted in the future. Nevertheless, current findings obtained from the macro-level perspective could provide some insights for people who are concerned about accident prevention of the entire industry.

Acknowledgements

The corresponding author, the long-term construction sector expert, was a full-time faculty at State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University. And the selection and weighting of indicators were supported by the experts about construction safety management from universities and construction companies. The authors would like to thank these supportive organizations and experts. Besides, the

authors would like to express the most sincere gratitude to the editors and reviewers for handling this manuscript and for the constructive comments and suggestions.

Funding

This work was supported by the National Natural Science Foundation of China under Grant No. 51779195; and the National Natural Science Foundation of China under Grant No. 51878385.

Author contributions

All of the authors were equally responsible for the conception and design of the study; Bo SHAO, Liyang TONG and Dawei LIU for data collection, methodology and paper writing; Bo SHAO, Zhigen HU and Xiazhong ZHENG for the literature and writing review.

Disclosure statement

The authors declare no conflict of interest.

References

- Bao, Q., Ruan, D., Shen, Y. J., Hermans, E., & Janssens, D. (2012). Improved hierarchical fuzzy TOPSIS for road safety performance evaluation. *Knowledge-Based Systems*, 32(8), 84-90. <https://doi.org/10.1016/j.knsys.2011.08.014>
- Bellamy, L. J. (2015). Exploring the relationship between major hazard, fatal and non-fatal accidents through outcomes and causes. *Safety Science*, 71, 93-103. <https://doi.org/10.1016/j.ssci.2014.02.009>
- Cai, W., Dou, L. M., Si, G. Y., Cao, A. Y., He, J., & Liu, S. (2016). A principal component analysis/fuzzy comprehensive evaluation model for coal burst liability assessment. *International Journal of Rock Mechanics and Mining Sciences*, 81, 62-69. <https://doi.org/10.1016/j.ijrmms.2015.09.028>
- Chen, S., Shao, B., Chen, Y., & Zheng, X. (2015). Dynamic risk classification of facilities in hydroelectric power station. *China Safety Science Journal*, 25(2), 71-76 (in Chinese). <http://doi.org/10.16265/j.cnki.issn1003-3033.2015.02.012>
- Coates, W. C. (2011). *Analysis of the causes of construction fatalities from 2007 to 2009*. Retrieved from <http://ufdc.ufl.edu/UFE0042920/00001>
- Cui, Y., Feng, P., Jin, J. L., & Liu, L. (2018). Water resources carrying capacity evaluation and diagnosis based on set pair analysis and improved the Entropy weight method. *Entropy*, 20(5), 359. <https://doi.org/10.3390/e20050359>
- Dai, S. Y., & Niu, D. X. (2017). Comprehensive evaluation of the sustainable development of power grid enterprises based on the model of fuzzy group ideal point method and combination weighting method with improved group order relation method and Entropy weight method. *Sustainability*, 9(10), 1900. <https://doi.org/10.3390/su9101900>
- Davoudabadi, R., Mousavi, S. M., Saparouskas, J., & Gitinavard, H. (2019). Solving construction project selection problem by a new uncertain weighting and ranking based on compromise solution with linear assignment approach. *Journal of Civil Engineering and Management*, 25(3), 241-251. <https://doi.org/10.3846/jcem.2019.8656>
- Depaire, B., Wets, G., & Vanhoof, K. (2008). Traffic accident segmentation by means of latent class clustering. *Accident Analysis and Prevention*, 40(4), 1257-1266. <https://doi.org/10.1016/j.aap.2008.01.007>
- Ding, L., Chen, K. L., Cheng, S. G., & Wang, X. (2015). Water ecological carrying capacity of urban lakes in the context of rapid urbanization: A case study of East Lake in Wuhan. *Physics and Chemistry of the Earth*, 89-90, 104-113. <https://doi.org/10.1016/j.pce.2015.08.004>
- Dong, X. S., Wang, X., Largay, J. A., Platner, J. W., Stafford, E., Cain, C. T., & Choi, S. D. (2015). Fatal falls in the U.S. residential construction industry. *American Journal of Industrial Medicine*, 57(9), 992-1000. <https://doi.org/10.1002/ajim.22341>
- Dong, X. W. S., Fujimoto, A., Ringen, K., Stafford, E., Platner, J. W., Gittleman, J. L., & Wang, X. W. (2011). Injury underreporting among small establishments in the construction industry. *American Journal of Industrial Medicine*, 54(5), 339-349. <https://doi.org/10.1002/ajim.20928>
- Dong, G. H., Shen, J. Q., Jia, Y. Z., & Sun, F. H. (2018). Comprehensive evaluation of water resource security: Case study from Luoyang City, China. *Water*, 10(8), 1106. <https://doi.org/10.3390/w10081106>
- Eteifa, S. O., & El-adaway, I. H. (2018). Using social network analysis to model the interaction between root causes of fatalities in the construction industry. *Journal of Management in Engineering*, 34(1), 1-15. [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000567](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000567)
- Fahad, A., Alshatri, N., Tari, Z., Alamri, A., Khalil, I., Zomaya, A. Y., Foufou, S., & Bouras, A. (2014). A survey of clustering algorithms for big data: Taxonomy and empirical analysis. *IEEE Transactions on Emerging Topics in Computing*, 2(3), 267-279. <https://doi.org/10.1109/TETC.2014.2330519>
- Fang, W. L., Ding, L. Y., Luo, H. B., & Love, P. E. D. (2018). Falls from heights: A computer vision-based approach for safety harness detection. *Automation in Construction*, 91, 53-61. <https://doi.org/10.1016/j.autcon.2018.02.018>
- Forteza, F. J., Carretero-Gomez, J. M., & Sese, A. (2017). Occupational risks, accidents on sites and economic performance of construction firms. *Safety Science*, 94, 61-76. <https://doi.org/10.1016/j.ssci.2017.01.003>
- Hasanzadeh, S., Esmaili, B., & Dodd, M. D. (2017). Impact of construction workers' hazard identification skills on their visual attention. *Journal Of Construction Engineering and Management*, 143(10), 04017070. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001373](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001373)
- Hausken, K., & Zhuang, J. (2016). The strategic interaction between a company and the government surrounding disasters. *Annals of Operations Research*, 237(1-2), 27-40. <https://doi.org/10.1007/s10479-014-1684-5>
- Hola, B. (2009). Methodology of estimation of accident situation in building industry. *Archives of Civil and Mechanical Engineering*, 9(1), 29-46. [https://doi.org/10.1016/S1644-9665\(12\)60038-7](https://doi.org/10.1016/S1644-9665(12)60038-7)
- Hola, B., & Nowobilski, T. (2018). Classification of economic regions with regards to selected factors characterizing the construction industry. *Sustainability*, 10(5), 1637. <https://doi.org/10.3390/su10051637>
- Hola, B., & Szostak, M. (2015). Analysis of the state of the accident rate in the construction industry in European Union countries. *Archives of Civil Engineering*, 61(4), 19-34. <https://doi.org/10.1515/ace-2015-0033>
- Horta, I. M., & Camanho, A. S. (2014). Competitive positioning and performance assessment in the construction industry. *Expert Systems with Applications*, 41(4), 974-983. <https://doi.org/10.1016/j.eswa.2013.06.064>

- Irumba, R. (2014). Spatial analysis of construction accidents in Kampala, Uganda. *Safety Science*, 64, 109–120. <https://doi.org/10.1016/j.ssci.2013.11.024>
- Jain, A. K. (2008). *Data clustering: 50 years beyond K-means*. Berlin, Heidelberg: Springer.
- Kang, Y., Siddiqui, S., Suk, S. J., Chi, S., & Kim, C. (2017). Trends of fall accidents in the US construction industry. *Journal of Construction Engineering and Management*, 143(8), 04017043. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001332](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001332)
- Karakhan, A. A., Rajendran, S., Gambatese, J., & Nnaji, C. (2018). Measuring and evaluating safety maturity of construction contractors: Multicriteria decision-making approach. *Journal of Construction Engineering and Management*, 144(7), 04018054. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001503](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001503)
- Kim, E., Yu, I., Kim, K., & Kim, K. (2011). Optimal set of safety education considering individual characteristics of construction workers. *Canadian Journal of Civil Engineering*, 38(5), 506–518. <https://doi.org/10.1139/l11-024>
- Lai, C. G., Chen, X. H., Chen, X. Y., Wang, Z. L., Wu, X. S., & Zhao, S. W. (2015). A fuzzy comprehensive evaluation model for flood risk based on the combination weight of game theory. *Natural Hazards*, 77(2), 1243–1259. <https://doi.org/10.1007/s11069-015-1645-6>
- Li, X. X., Wang, K. S., Liu, L. W., Xin, J., Yang, H. R., & Gao, C. Y. (2011). Application of the Entropy weight and TOPSIS method in safety evaluation of coal mines. *Procedia Engineering*, 26(4), 2085–2091. <https://doi.org/10.1016/j.proeng.2011.11.2410>
- Li, G., Li, J., Sun, X., & Zhao, M. (2017). Research on a combined method of subjective-objective weighing and the its rationality. *Management Review*, 29(12), 17–26. <http://doi.org/10.14120/j.cnki.cn11-5057/f.2017.12.002>
- Lin, Y. (1989). *Numerical classification method and its application in rock mechanics*. Shenyang: Northeast University Press.
- Liu, A., & Wu, C. (2011). Correlation analysis between death rate per hundred million GDP and regional development level. *China Safety Science Journal*, 21(5), 3–9 (in Chinese). <http://doi.org/10.16265/j.cnki.issn1003-3033.2011.05.019>
- Liu, F., Zhao, S. Z., Weng, M. C., & Liu, Y. Q. (2017). Fire risk assessment for large-scale commercial buildings based on structure Entropy weight method. *Safety Science*, 94, 26–40. <https://doi.org/10.1016/j.ssci.2016.12.009>
- Ma, J., Chen, X., & Liu, W. (2015). Suggestions on the supervision and assessment of safety production in the construction industry. *Construction and Architecture*, 17, 34–37 (in Chinese). <http://doi.org/10.3969/j.issn.0577-7429.2015.17.012>
- Ma, Y. H., & Zhao, Q. H. (2018). Decision-making in safety efforts: Role of the government in reducing the probability of workplace accidents in China. *Safety Science*, 104, 81–90. <https://doi.org/10.1016/j.ssci.2017.12.038>
- Mahmoudzadeh, M., & Bafandeh, A. R. (2013). A new method for consistency test in fuzzy AHP. *Journal of Intelligent & Fuzzy Systems*, 25(2), 457–461. <http://doi.org/10.3233/ifs-120653>
- Marsh, S. M., & Fosbroke, D. E. (2015). Trends of occupational fatalities involving machines, United States, 1992–2010. *American Journal of Industrial Medicine*, 58(11), 1160–1173. <https://doi.org/10.1002/ajim.22532>
- Mendeloff, J., & Burns, R. (2013). States with low non-fatal injury rates have high fatality rates and vice-versa. *American Journal of Industrial Medicine*, 56(5), 509–519. <https://doi.org/10.1002/ajim.22047>
- Ministry of Housing and Urban-Rural Development of China (MHURD). (2019a). *Announcement on the accident situation of production safety of the China housing and municipal engineering in 2018*. Retrieved from http://www.mohurd.gov.cn/wjfb/201903/t20190326_239913.html
- Ministry of Housing and Urban-Rural Development of China (MHURD). (2019b). *The short reports of fatal accidents in China's building construction activities*. Retrieved from <http://sgxxxt.mohurd.gov.cn/Public/AccidentList.aspx>
- National Bureau of Statistics of China (NBS). (2019). *National data*. Retrieved from <http://data.stats.gov.cn/easyquery.htm?cn=C01>
- Poh, C. Q. X., Ubeynarayana, C. U., & Goh, Y. M. (2018). Safety leading indicators for construction sites: A machine learning approach. *Automation in Construction*, 93, 375–386. <https://doi.org/10.1016/j.autcon.2018.03.022>
- Prascevic, N., & Prascevic, Z. (2017). Application of fuzzy AHP for ranking and selection of alternatives in construction project management. *Journal of Civil Engineering and Management*, 23(8), 1123–1135. <https://doi.org/10.3846/13923730.2017.1388278>
- Raviv, G., Fishbain, B., & Shapira, A. (2017). Analyzing risk factors in crane-related near-miss and accident reports. *Safety Science*, 91, 192–205. <https://doi.org/10.1016/j.ssci.2016.08.022>
- Sari, M., Selcuk, A. S., Karpuz, C., & Duzgun, H. S. B. (2009). Stochastic modeling of accident risks associated with an underground coal mine in Turkey. *Safety Science*, 47(1), 78–87. <https://doi.org/10.1016/j.ssci.2007.12.004>
- Shao, B., Hu, Z., Liu, Q., Chen, S., & He, W. (2019). Fatal accident patterns of building construction activities in China. *Safety Science*, 111, 253–263. <https://doi.org/10.1016/j.ssci.2018.07.019>
- Soltanzadeh, A., Mohammadfam, I., Moghimbeigi, A., & Ghiasvand, R. (2017). Exploring causal factors on the severity rate of occupational accidents in construction worksites. *International Journal of Civil Engineering*, 15(7A), 959–965. <https://doi.org/10.1007/s40999-017-0184-9>
- Sunindijo, R. Y., & Zou, P. X. W. (2012). Political skill for developing construction safety climate. *Journal of Construction Engineering and Management*, 138(5), 605–612. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000482](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000482)
- Tam, C. M., Zeng, S. X., & Deng, Z. M. (2004). Identifying elements of poor construction safety management in China. *Safety Science*, 42(7), 569–586. <https://doi.org/10.1016/j.ssci.2003.09.001>
- Wang, B., Wu, C., Huang, L., Zhang, L., Kang, L., & Gao, K. (2018a). Prevention and control of major accidents (MAs) and particularly serious accidents (PSAs) in the industrial domain in China: Current status, recent efforts and future prospects. *Process Safety and Environmental Protection*, 117, 254–266. <https://doi.org/10.1016/j.psep.2018.04.025>
- Wang, B., Wu, C., Kang, L., Huang, L., & Pan, W. (2018b). What are the new challenges, goals, and tasks of occupational health in China's Thirteenth Five-Year Plan (13th FYP) period? *Journal of Occupational Health*, 60(3), 208–228. <https://doi.org/10.1539/joh.2017-0275-RA>
- Wang, B., Wu, C., Kang, L., Reniers, G., & Huang, L. (2018c). Work safety in China's Thirteenth Five-Year Plan period (2016–2020): Current status, new challenges and future tasks. *Safety Science*, 104, 164–178. <https://doi.org/10.1016/j.ssci.2018.01.012>
- Wang, B., Wu, C., Reniers, G., Huang, L., Kang, L., & Zhang, L. (2018d). The future of hazardous chemical safety in China: Opportunities, problems, challenges and tasks. *Science of the Total Environment*, 643, 1–11. <https://doi.org/10.1016/j.scitotenv.2018.06.174>

- Wegman, F., & Oppe, S. (2010). Benchmarking road safety performances of countries. *Safety Science*, 48(9), 1203-1211. <https://doi.org/10.1016/j.ssci.2010.02.003>
- Wu, X., Liu, Q., Zhang, L., Skibniewski, M. J., & Wang, Y. (2015). Prospective safety performance evaluation on construction sites. *Accident Analysis and Prevention*, 78, 58-72. <https://doi.org/10.1016/j.aap.2015.02.003>
- Wu, D. F., Wang, N. L., Yang, Z. P., Li, C. Z., & Yang, Y. P. (2018). Comprehensive evaluation of coal-fired power units using grey relational analysis and a hybrid Entropy-based weighting method. *Entropy*, 20(4), 215. <https://doi.org/10.3390/e20040215>
- Xiahou, X., Yuan, J. F., Li, Q. M., & Skibniewski, M. J. (2018). Validating DFS concept in lifecycle subway projects in China based on incident case analysis and network analysis. *Journal of Civil Engineering and Management*, 24(1), 53-66. <https://doi.org/10.3846/jcem.2018.300>
- Yassin, A. S., & Martonik, J. F. (2004). The effectiveness of the revised scaffold safety standard in the construction industry. *Safety Science*, 42(10), 921-931. <https://doi.org/10.1016/j.ssci.2004.05.001>
- Yuan, Z. (2005). Is the construction safety good or poor? *Construction Enterprise Management*, 5(5), 46-47 (in Chinese). <http://doi.org/10.3969/j.issn.1001-9251.2005.05.024>
- Zhang, W., Zhang, X., Luo, X. W., & Zhao, T. S. (2019). Reliability model and critical factors identification of construction safety management based on system thinking. *Journal of Civil Engineering and Management*, 25(4), 362-379. <https://doi.org/10.3846/jcem.2019.8652>
- Zhao, X., Guo, H. T., Huang, C. L., & Zhong, J. S. (2017). Teaching evaluation system research based on structure entropy weight method. *Journal of Discrete Mathematical Sciences & Cryptography*, 20(1), 179-191. <https://doi.org/10.1080/09720529.2016.1178915>
- Zheng, X., Shao, B., Chen, L., Chen, S., & Ge, S. (2014). Safety entropy evaluation of hydropower construction based on Euclid theory. *China Safety Science Journal*, 24(6), 38-43 (in Chinese). <http://doi.org/10.16265/j.cnki.issn1003-3033.2014.06.013>
- Zheng, X., Zhou, J., Wang, F., & Chen, Y. (2018). Routes to failure and prevention recommendations in work systems of hydropower construction. *Journal of Civil Engineering and Management*, 24(3), 206-222. <https://doi.org/10.3846/jcem.2018.1647>
- Zhou, Z. P., Goh, Y. M., & Li, Q. M. (2015). Overview and analysis of safety management studies in the construction industry. *Safety Science*, 72, 337-350. <https://doi.org/10.1016/j.ssci.2014.10.006>