Safety Science 50 (2012) 1019–1033

Contents lists available at SciVerse ScienceDirect

Safety Science

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

Fatality and influence factors in high-casualty fires: A correspondence analysis

Song Lu, Peng Mei, Jie Wang, Heping Zhang *

State Key Laboratory of Fire Science, University of Science and Technology of China, Hefei, Anhui 230026, China

ARTICLE INFO

Article history:
Received 28 April 2011
Received in revised form 11 November 2011
Accepted 5 December 2011
Available online 27 December 2011

Keywords:
Correspondence analysis
Fatality statistics
Influence factor
High-causality fires

ABSTRACT

The technology of correspondence analysis was applied to high-casualty fire data in China. The aim of this study was to investigate the associations between fatality levels and influence factors that involve place, cause, time of day, month, year and province. The variable fatality in a fire has four levels: 3, 4–5, 6–9 and $\geqslant 10$. The results show that hotels, welfare houses and hospitals tend to be strongly associated with fatality level $\geqslant 10$. The fires caused by work-related tasks tend to precipitate a relatively high number of fatalities and are strongly associated with fatality level 6–9. Fires that occur in the daytime (8:00–19:59) are associated with higher fatalities than fires that occur at night (20:00–7:59). The months in the cold season, such as winter or the beginning of spring, tend to be associated with fatality levels 4–5, 6–9 and $\geqslant 10$. CA dynamically portrayed the fatality tendency over the past 8 years, and 2007 tended to be associated with fatality level $\geqslant 10$. Fatality characteristics of provinces are identified, and Beijing, Shandong and Jilin are strongly associated with fatality level $\geqslant 10$. To explore whether associations between influence factors and fatality levels of high-casualty fires in China resemble corresponding associations of HCFs in the United States, data on fires with fatality level $\geqslant 5$ in the two countries were collected. The results of four sets of comparisons indicate that the associations between influences and fatality levels in the two countries present contrasting features. Some practical applications are briefly discussed.

1. Introduction

Recently, a succession of high-casualty fires (HCFs) have occurred in China: for example, the November 5, 2010, Jilin shopping-mall fire with a death toll of 19; the November 15, 2010, Shanghai high-rise fire, with a death toll of 59; and the January 18, 2011, Wuhan fire, with a death toll of 14. Because life is a priceless treasure, a high number of casualties result in a great deal of human suffering. To lower the risk of HCFs, it is necessary to study them carefully.

The relationships between fire casualties and their relevant influence factors have been studied for many years and still present meaningful challenges. Early studies in the relationships between fire fatalities and influence factors, such as year, season, region, occupancy and fire company response distances, were reported in 1976 by Corman et al. (1976). The associations between fatalities and population, location, time and behavior were investigated by Barillo and Goode (1996). Jordan et al. (1999) compared the influence of important factors in fire fatalities over a 10-year period and identified changes in fatality trends. Holborn et al. (2003) systematically identified the main factors that were involved in the deaths of individuals in non-arson residential fires. The work conducted by Duncanson et al. (2002) focused on the relationship between socioeconomic deprivation and the risk of a non-arson fatal domestic fire incident. Hasofer and Thomas (2006) explored the fires and the personal characteristics that result in the greatest risk of casualties in 2006 using generalized linear models and analysis of deviance. There have also been many studies that focused only on fatalities that involve one type of individual. For example, Elder et al. (1996) provided the characteristics of elderly individuals who died in residential fires between 1980 and 1990, and Squires and Busuttil (1995) focused on the identifying factors that precipitated child fatalities in Scottish house fires.

The problems associated with HCFs in China have, for a long time, been a subject of great interest to Chinese researchers. Yang et al. (2002) and Guo and Fu (2007) reported and analyzed the situations of HCFs in China based on the relevant statistics, and they pointed out that HCFs had become a noticeable problem. Attempts have been made to investigate the relationships between influence factors and fire situations. The effect of socioeconomic factors on fire in China was discussed by Yang et al. (2005), whereas Chen et al. (2007) performed cluster analysis on fire statistics (2000–2002) and discussed the relationship between the development of economy, fire protection devotion and fire loss. Furthermore, Li (2010) analyzed serious fire disasters (with fatalities $\geqslant 30$) in China (2000–2008) and qualitatively discussed the extent of influence of several risk factors.

The associations between factors and fatalities have been analyzed and have formed the cornerstone of most fire prevention strategies. HCFs involve different fatality levels. From a literature

---

* Corresponding author. Tel.: +86 551 3601665; fax: +86 551 3606981.
E-mail address: zhanghp@ustc.edu.cn (H. Zhang).

0925-7535/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ssci.2011.12.006

survey, it appears that none of the previous investigations were concerned with associations between factors and different fatality levels. In previous work, most analyses were within the scope of the total number of fatalities, irrespective of various fatality levels. For example, suppose that we are interested in the following two questions: What are the similarities and differences among the eight fire causes with respect to the four fatality levels? And what are the associations among the eight fire causes and four fatality levels? A proper understanding of the associations between factors and different fatality levels may be at the heart of effective fire prevention strategies and the accurate quantification of fire risks.

Compared with most other branches of science and technology, wherein experiments offer the basis for building, testing and refining theoretical frameworks, the ability to conduct such experiments in the domain of accident forecasting and prevention is limited. Among the various methodologies of fire studies, statistical analysis has proven advantageous. Correspondence analysis (CA) is a specialized method of multivariate statistical analysis (Johnson and Wichern, 2007). CA has two merits that contribute to its usefulness for exploring the associations between factors and different fatality levels. First, CA applies multivariate analysis to the data through the simultaneous consideration of multiple categorical variables. The multivariate nature of CA can reveal relationships that would not be detected in a series of pairwise comparisons of variables and can be investigated by the multivariate features of CA (Hoffman and Franke, 1986). Second, CA helps to portray how variables are associated, not merely that an association exists. The correspondence analysis map contributes to investigating structural relationships among the variable categories, whereas a tabular approach is less effective because of its large tables. Although CA has been applied in depth in many fields (Sourial et al., 2010; Park et al., 2007; Pusha et al., 2009; Higgs, 1990), it has not yet been extensively used in fire research.

The main research motivation of the present work is fourfold. The first research motivation is to investigate the associations between influence factors and different fatality levels. The factors studied were as follows: place, cause, time of day, month, year and province. The associations between a certain influence factor and different fatality levels might be significantly different. Although we can easily obtain an association between an influence factor and all of the fires with fatalities, the result is not very helpful when developing effective fire prevention strategies. The reason is that the number of fires with relatively low fatalities is much larger than the number of high fatality fires, and this difference in numbers may mask the characteristics of fires with very high fatalities. The second motivation is to enhance our understanding of a fire with a very high fatality level. The reason for the second motivation is that fires with very high casualties, for example, fatalities $\geqslant 10$, have received considerable attention in recent years, but less work has been performed on the corresponding fire statistics research. The third motivation is to explore whether relationships between influence factors and fatality levels of HCFs in China show similarities to the corresponding relationships of HCFs in the United States. These comparisons are interesting and meaningful for fire administration in the two countries. The fourth motivation is to demonstrate that CA can provide a new perspective on exploring the relationships between influence factors and casualty levels.

Applying CA to the HCF dataset can obtain the relationships between influence factors and fatality levels. In every factor, there are some objects that tend to be associated with fires that precipitate a very high fatality level. Because no study on the associations between specific objects and fatality levels of HCFs has been reported, these results make a contribution to the development of effective fire prevention strategies for controlling HCFs. To explore whether similar associations exist in American HCFs, a series of comparisons are conducted. It is found that the Chinese HCFs tend to be associated with higher fatalities than do HCFs in the United States; the relationships between the influencing factors and the fatality levels in the United States are quite unlike the corresponding relationships in China. Finally, the analysis results demonstrate the power of CA in exploring the relationships among categorical variables.

2. Data and method

2.1. Data

Given the difficulty of obtaining the number of injured victims, HCFs are represented by fires with three or more fatalities in the present study. For the purpose of analysis, the variable fatality in a fire has four levels, 3, 4–5, 6–9 and $\geqslant 10$, and these values represent the number of fatalities. Only 735 fire samples with more than three fatalities in mainland China from 2002 to 2009 are available. It should be noted that forest fires are excluded. Most fires were located in cities and rural areas, and only a few occurred in vehicles. In China, all of the fire statistics are compiled by the Fire Service Bureau and the Ministry of Public Security; therefore, the fire samples are derived from the Fire Statistical Year Book of China 2003 (Fire Service Bureau, 2003) and China Fire Services 2004–2010 (Fire Service Bureau, 2005a,b, 2006, 2007, 2008, 2009, 2010), which are both edited by the Fire Service Bureau and are the most official fire statistics that are available at the present time.

Some American catastrophic multiple-death fire data (2003–2009) (Badger, 2004, 2005, 2006, 2007, 2008, 2009, 2010) are used for conducting comparisons. In the United States, catastrophic multiple-death fire refers to either a fire or an explosion with a fire in a home or apartment with five or more fire-related deaths or a fire or explosion in any other structure, as well as outside of structures (such as wildfires and vehicle fires), that claims three or more lives. Catastrophic multiple-death fires belong to HCFs based on the definition of HCFs in the present work. Because the dataset of catastrophic multiple-death fires in the United States has some features that are not considered in China fire statistical work, the two datasets cannot be compared directly. To ensure the uniformity of two datasets, many fire samples are discarded in the comparison analysis. In the present work, only fire samples that have no less than five fatalities and that occurred in 2003–2009 are selected. There are 162 and 131 HCFs with $\geqslant 5$ fatalities in China and the United States, respectively. For the purpose of being consistent with previous analysis, the variable fatality in a fire is divided into three levels: 5, 6–9 and $\geqslant 10$. Because too many fire samples in the United States lack a cause of the fire, the CA of cause vs. fatality cannot be conducted in comparison. The CA of province vs. fatality is also not performed because the number of fire samples is so small that the contingency table for province vs. fatality is a sparse matrix. CA cannot conduct analysis with a very sparse contingency table.

Despite the inherent shortcomings of the fire statistics, which mainly arise from under-reporting or improper reporting of fire accidents in China, statistics remain important for fire prevention. However, these shortcomings undoubtedly influence the CA results. Because fires with more than three fatalities are serious fires that receive significant public and media attention, these records are more credible than small fires, and the probability of improper reporting is low. In contrast, the probability of under-reporting is more or less high. The main reason is that some organizations that are responsible for HCFs and that might be punished by governments tend to conceal the HCFs. From a literature survey, it appears that none of the previous investigations are concerned with the under-reporting of fire casualties. Many studies might have been hampered by the lack of a comparative dataset. According to Amoros et al. (2006), to establish the under-reporting rate, a comparative dataset is essential. With respect to fire casualties, the dataset from fire services and the dataset from an insurance company can serve as one comparative dataset. In China, the fire statistics work is exclusively performed by the fire service, and other companies and organizations are not permitted to directly investigate and collect fire casualty information. Hence, we cannot acquire a valuable comparative dataset and conduct a reasonable analysis of the influence due to under-reporting fire casualties in the present work. Despite much effort, the objective of eliminating the influence of under-reporting often leaves much to be desired in the present work. Here, we discuss the factors that influence the under-reporting rate qualitatively. In common with other accidents, under-reporting of fire accidents is inversely and strongly associated with casualty severity. The under-reporting rate of HCFs might be relatively low because the casualties from HCFs are very serious. Geographic factors may influence the under-reporting rate as well. For example, fires that occur in very remote areas are easily under-reported because the fire service cannot effectively cover all of these remote regions. Time also can have an effect, to some extent, on the under-reporting rate. Because the processes and technologies for fire statistics work are increasingly standard and advanced with time, the under-reporting rate might have been reduced in recent years. Although this change is positive, it increases the difficulty of under-reporting analysis. The above three factors could be the three most important reasons that the impact on the fire under-reporting rate affects the results of CA. The limitations of under-reporting will hopefully be addressed in future research.

2.2. Method

For subsequent reference, the basic algorithm of CA is discussed first. A more complete discussion of CA is reported by Clausen (1998). In the following discussion, boldface capital letters denote matrices, boldface lowercase letters denote vectors, and lowercase italic letters denote scalars. The aim of CA is to find a low-rank approximation of a contingency table, which is a two-way frequency table where the joint frequencies of two categorical variables are recorded. Let **X** denote the $n \times p$ rows-by-columns categorical data matrix displayed in the contingency table.

Each row of **X** represents a point profile in p-dimensional space, and each column represents a point profile in n-dimensional space. As the raw frequencies in the contingency table do not yield a meaningful interpretation of distances between row points and between column points, the profiles of raw occurrences should convert to their frequency distributions. CA begins by transforming the frequencies in a classification into fractions.

$$\mathbf{P} = \frac{\mathbf{X}}{\mathbf{1}'\mathbf{X}\mathbf{1}}, \quad \textbf{with} \quad \mathbf{1}'\mathbf{P}\mathbf{1} = 1 \tag{1}$$

where $\mathbf{1}' = (1, \ldots, 1)'$, either $1 \times n$ or $1 \times p$, relying on the context. **P** is the probability density of the cells of **X**. For the sake of calculation, we define $\mathbf{D}_r$ and $\mathbf{D}_c$. The row sums of **P** are written into $\mathbf{D}_r$, an $n \times n$ diagonal matrix,

$$\mathbf{D}_r = diag(\mathbf{r}) \tag{2}$$

where **r** = **P1**, and the column sums of **P** are written into $\mathbf{D}_c$, a $p \times p$ matrix,

$$\mathbf{D}_c = diag(\mathbf{c}) \tag{3}$$

where **c** = **P′1**. In CA, **r** and **c** are referred to as masses of row and column points, respectively. In the present work, we are working with a contingency table, and **r** and **c** are the marginal densities.

CA is a graphical procedure for representing associations; it graphically represents the distance between row (or column) profiles. The configuration of points at the "center of gravity" of both sets is oriented. The centroid of the set of row variables in its space is **c**, the vector of column masses. The centroid of the set of column variables in its space is **r**, the vector of row masses. To perform the analysis relative to the center of gravity, **P** is centered symmetrically by rows and columns, i.e., $\mathbf{P} - \mathbf{rc}'$; thus, the origin corresponds to the average profile of both sets of points.

The next step is to compute the singular value decomposition (SVD) of $\mathbf{D}_r^{-1/2}(\mathbf{P} - \mathbf{rc}')\mathbf{D}_c^{-1/2}$, as follows:

$$\mathbf{D}_r^{-1/2}(\mathbf{P} - \mathbf{rc}')\mathbf{D}_c^{-1/2} = \mathbf{M}\mathbf{D}_\mu\mathbf{N}' \tag{4}$$

where **M** ($n \times k$) and **N** ($p \times k$) are diagonal matrices, **M′M** = **N′N** = **I**, and $\mathbf{D}_\mu$ ($k \times k$) is a diagonal matrix of singular values $\mu_1, \cdots, \mu_t, \cdots \mu_k$, with $\mu_1 \geqslant \ldots \geqslant \mu_t \geqslant \ldots \geqslant \mu_k > 0$. Then, the coordinates of the row and column points in k dimensions are given by the rows of **F** and **G**, respectively:

$$\mathbf{F} = \mathbf{D}_r^{-1/2}\mathbf{M}\mathbf{D}_\mu \tag{5}$$

and

$$\mathbf{G} = \mathbf{D}_c^{-1/2}\mathbf{N}\mathbf{D}_\mu \tag{6}$$

According to **F** and **G**, the points are plotted to visualize the associations among the variables. This graphical representation is called a CA map.

Once a CA map is established, we can interpret it and find associations among row and column points. By referring to Johnson and Wichern (2007), row points that are close together indicate rows that have similar profiles across the columns; column points that are close together indicate columns that have similar profiles across the rows. Because the distances are defined only within rows (mortality intervals) or within columns (fire causes), not across the rows and columns, it is not appropriate to directly compare a row and a column point. Hoffman and Franke (1986) suggested that if the dimensions (or principal axes) are initially defined, then the relative positioning of the row and column points could be defined within those dimensions. The associations between the row points and the column points can be assessed by their relative positions within the dimensions.

To interpret the CA map precisely, the following variables are defined. Inertia is defined as the weighted sum of the squared chi-squared distance between each profile and the average profile. The inertia refers to a degree of variance or dispersion of the individual profiles around the average profile. The overall spatial variation in the row and column points can be quantified and are conducive to the interpretation of the CA results. The total inertia is defined as the weighted sum of squared distances from the points to their respective centroids and is equivalent for both row and column points.

$$Inertia(Total) = \sum_i \sum_j \frac{(p_{ij} - r_i c_j)^2}{r_i c_j} \tag{7}$$

The total inertia can be decomposed along the principal axes and written as the sum of eigenvalues.

$$Inertia(Total) = \sum_{t=1}^{k} \mu_t^2 = \sum_{t=1}^{k} \lambda_t \tag{8}$$

where each eigenvalue λ_t indicates the weighted variance (inertia) explained by the tth principal axis.

The inertia of the ith row point is defined as:

$$r_i \sum_t f_{it}^2 \tag{9}$$

where f_{it} is the element of the matrix **F**, and r_i is the mass of the ith row point. The above equation represents the contribution of the ith row point to the total inertia. A similar definition holds for each column

Table 1
Contingency table for place and fatality.

Place	Fatality				Total
	3	4–5	6–9	⩾10	
Dwelling	180	78	15	3	276
Shop	88	52	13	2	155
Mall and marketplace	9	9	7	4	29
Family workshop	12	6	9	4	31
Industrial and storage	21	18	10	7	56
Construction site	4	4	2	1	11
Welfare house and hospital	2	4	6	1	13
Restaurant and entertainment	22	15	5	8	50
Hotel	4	4	4	4	16
Dormitory	4	5	3	1	13
Others	47	21	12	5	85
Total	393	216	86	40	735

point. These contributions, summed over all of the row points (or column points), equal the total inertia.

The inertia along the tth axis, λ_t, consists of the weighted sum of squared distances to the origin of the displayed row (or column) point profiles. For row point profiles, this inertia can be defined as:

$$\lambda_t = \sum_i r_i f_{it}^2 \tag{10}$$

A similar definition holds for the column point profiles. Therefore, the inertia of the projections of the row points (or column points) on each axis are represented by the corresponding eigenvalues.

The absolute contribution of the ith row point to the tth principal axis is defined as:

$$\frac{r_i f_{it}^2}{\lambda_t} \tag{11}$$

The absolute contributions measure the importance of each point in determining the direction of the principal axes, which serve as guides to the interpretation of the associations between the row and column points.

The relative contribution to inertia measures the quality of the representation of each point in the display. The relative contribution of the tth principal axis to the inertia of the ith row point is written as:

$$\frac{f_{it}^2}{\Sigma_t f_{it}^2} \tag{12}$$

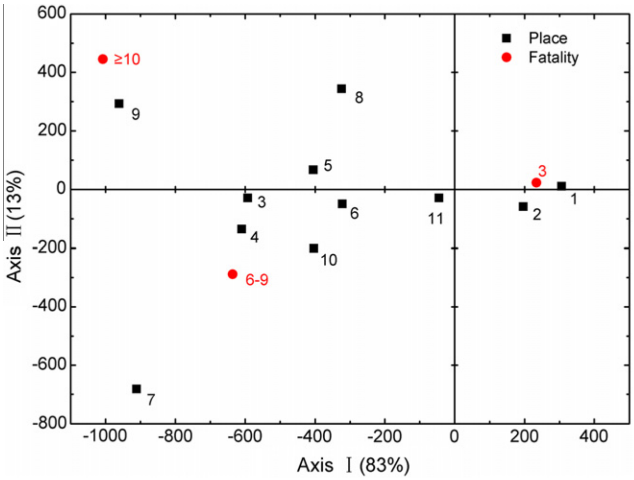

Fig. 1. CA map of place vs. fatality. Place: 1. Dwelling, 2. Shop, 3. Mall and marketplace, 4. Family workshop, 5. Industrial and storage, 6. Construction site, 7. Welfare house and hospital, 8. Restaurant and entertainment, 9. Hotel, 10. Dormitory, 11. Others.

Table 3
Contingency table for the cause of the fire and fatality.

Cause	Fatality				Total
	3	4–5	6–9	⩾10	
Electricity	123	70	37	15	245
Improperly using fire in daily life	113	51	15	5	184
Work-related tasks	30	17	7	7	61
Arson	48	35	11	7	101
Playing with fire	27	7	5	3	42
Smoking	10	4	3	1	18
Unknown	30	17	4	1	52
Others	12	15	4	1	32
Total	393	216	86	40	735

A relative contribution is equal to the $\cos^2$ of the angle θ between the point and the tth principal axis and is a square correlation. An analysis should focus on the high-quality points, and less focus should be placed on points that are not well explained by principal axes (Garson, 2008).

Table 2
Decomposition of inertia among the place and fatality for the first two principles[a].

Object	Quality	Mass	Inertia	Axis I			Axis II		
				Coordinate	Squared correlation	Contribution	Coordinate	Squared correlation	Contribution
Place									
Dwelling	989	376	220	306	988	266	11	1	2
Shop	857	211	55	196	786	62	−59	71	35
Mall and marketplace	993	39	87	−593	990	105	−29	2	2
Family workshop	898	42	103	−610	857	119	−135	42	37
Industrial and storage	969	76	81	−405	943	95	67	26	16
Construction site	806	15	10	−322	788	12	−49	18	2
Welfare house and hospital	999	18	144	−911	640	111	−682	359	394
Restaurant and entertainment	994	68	95	−324	468	54	344	527	384
Hotel	997	22	138	−961	912	152	293	85	90
Dormitory	833	18	23	−404	668	22	−201	165	34
Others	211	116	2	−45	149	2	−29	62	5
Fatality									
3	950	535	185	234	941	221	23	9	14
4–5	19	294	1	15	13	0	−9	5	1
6–9	994	117	359	−636	823	359	−289	170	469
⩾10	999	54	414	−1007	836	419	445	163	516

[a] All values are multiplied by 1000 and decimal points are omitted.

Table 4
Decomposition of inertia among the cause and fatality for the first two principles[a].

Object	Quality	Mass	Inertia	Axis I			Axis II		
				Coordinate	Squared correlation	Contribution	Coordinate	Squared correlation	Contribution
Cause									
Electricity	654	333	74	87	584	130	30	71	21
Improperly using fire in daily life	994	250	242	−192	988	473	−15	6	4
Work-related tasks	694	83	107	185	480	146	123	214	89
Arson	774	137	58	99	475	70	−79	299	60
Playing with fire	1000	57	123	−94	108	26	271	891	294
Smoking	546	24	14	23	13	1	145	533	36
Unknown	992	71	80	−167	644	102	−123	348	75
Others	994	44	184	154	146	53	−371	848	420
Fatality									
3	998	535	220	−112	802	347	55	196	116
4–5	986	294	259	60	107	55	−173	879	620
6–9	627	117	125	189	553	216	69	74	39
⩾10	874	54	278	369	611	382	242	263	225

[a] All values are multiplied by 1000 and decimal points are omitted.

Table 5
Contingency table for the time of day and fatality.

Time of day	Fatality				Total
	3	4–5	6–9	⩾10	
0:00–3:59	153	92	32	10	287
4:00–7:59	67	50	18	5	140
8:00–11:59	29	11	8	5	53
12:00–15:59	35	11	5	6	57
16:00–19:59	28	13	6	7	54
20:00–23:59	70	30	14	7	121
Total	382	207	83	40	712

Table 7
Contingency table for the month of the fire and fatality.

Month	Fatality				Total
	3	4–5	6–9	⩾10	
January	56	37	10	3	106
February	41	25	8	8	82
March	25	21	10	3	59
April	33	11	4	1	49
May	38	16	7	3	64
June	27	11	6	3	47
July	32	12	2	1	47
August	22	7	7	1	37
September	26	15	4	4	49
October	26	16	7	4	53
November	27	17	9	4	57
December	40	28	12	5	85
Total	393	216	86	40	735

3. HCFs in China

3.1. Place vs. fatality

The contingency table of place and fatality is shown in Table 1. The eigenvalues for the first two-dimensional axes are 0.363 and 0.145, with cumulative proportions of inertia equaling 83% and 96%, respectively. Typically, two-dimensional displays are often satisfactory if the cumulative inertia is sufficiently large, and commonly used rules recommend that the two dimensions retained represent >70% inertia in market research fields (Higgs, 1990). In the present study, the criterion of cumulative inertia >70% is adopted. The two-dimensional map and numerical results from a CA of the type of place vs. fatality are shown in Fig. 1 and Table 2, respectively. In Table 2, the absolute contribution and relative contribution are presented in the columns entitled “Contribution” and “Squared correlation,” respectively; the relative contributions of each point in the first two-dimensional space are titled “Quality.” These labels are the same as those shown in Tables 4, 6, 8, 10 and 12.

First, we assess the similarity between the fatalities and the places. The fatality levels 3, 6–9 and ⩾10 have very different place profiles, and the place profile of the fatality level 4–5 is similar to the average profile. In terms of places, several patterns emerge. Dwellings and shops have similar fatality profiles; malls and marketplaces, family workshops, industrial and storage facilities,

Table 6
Decomposition of inertia among the time of day and fatality for the first two principles[a].

Object	Quality	Mass	Inertia	Axis I			Axis II		
				Coordinate	Squared correlation	Contribution	Coordinate	Squared correlation	Contribution
Time of day									
0:00–3:59	985	403	141	103	932	160	25	53	62
4:00–7:59	999	197	184	158	827	183	−72	172	261
8:00–11:59	795	74	110	−215	769	129	−39	26	30
12:00–15:59	967	80	227	−294	918	259	68	49	94
16:00–19:59	975	76	242	−292	808	242	−133	167	342
20:00–23:59	929	170	47	−64	428	26	70	502	211
Fatality									
3	999	537	86	−46	404	42	55	595	421
4–5	975	291	314	179	901	349	−51	74	197
6–9	280	117	15	22	34	2	−59	246	106
⩾10	995	56	538	−537	933	607	−139	62	277

[a] All values are multiplied by 1000 and decimal points are omitted.

Table 8
Decomposition of inertia among the month and fatality for the first two principles.[a]

Object	Quality	Mass	Inertia	Axis I			Axis II		
				Coordinate	Squared correlation	Contribution	Coordinate	Squared correlation	Contribution
Month									
January	107	144	10	−40	59	10	−36	48	19
February	752	112	83	99	246	48	−142	506	228
March	827	80	99	205	703	149	86	124	60
April	998	67	139	−288	982	243	37	16	9
May	856	87	27	−108	801	45	28	55	7
June	224	64	6	−43	103	5	46	121	14
July	989	64	177	−325	937	297	−77	52	38
August	923	50	118	−82	66	15	296	857	445
September	884	67	36	19	15	1	−147	869	145
October	849	72	21	108	834	37	−14	15	2
November	943	78	47	145	810	72	59	134	27
December	903	116	45	123	874	77	22	29	6
Fatality									
3	964	535	236	−133	963	418	4	1	1
4–5	562	294	109	104	407	140	−64	155	122
6–9	991	117	283	206	430	218	235	561	651
⩾10	659	54	182	306	458	224	−203	201	226

[a] All values are multiplied by 1000 and decimal points are omitted.

Table 9
Contingency table for the year and fatality.

Year	Fatality				Total
	3	4–5	6–9	⩾10	
2002	69	40	11	3	123
2003	58	35	14	6	113
2004	54	38	9	4	105
2005	54	24	10	7	95
2006	52	14	14	3	83
2007	33	22	10	9	74
2008	39	25	11	4	79
2009	34	18	7	4	63
Total	393	216	86	40	735

construction sites and dormitories have similar fatality profiles as well. However, the fatality profiles of welfare houses, hospitals, restaurants, places of entertainment and hotels are different. As the “Quality” of fatality point 4–5 (quality = 0.019) is very low, indicating that it has the worst fit in the plane defined by the first two principal axes, less analytical focus must be placed on this point.

According to the contributions in Table 2, fatality level ⩾10 is the primary contributor to axis I, followed by fatality levels 6–9 and 3. Note that fatality levels ⩾10 and 3 are extreme in terms of their location on axis I (for example, the highest or lowest values on axis I). Meanwhile, fatality levels ⩾10 and 6–9 are relatively close in terms of their location on axis I. Therefore, the direction of axis I can be defined by fatality points, and axis I separates low fatality points on the right from high fatality points on the left. With an increase in the abscissa, fatalities decrease. We label axis I as a “fatality” axis, which plays a decisive role in the analysis that follows. Axis II cannot be meaningfully labeled as in marketing research (Hoffman and Franke, 1986) because fatalities cannot be reasonably explained as “brands” characterized by potential attributes. After the “fatality” dimension and its direction are defined, the association between fatality and place can be analyzed.

In Fig. 1, dwelling and shop place points are on the low fatality side, which indicates that they are places that tend to be associated with fires that cause a relatively low number of fatalities. In dwellings and shops, the numbers of occupants are fewer compared with other places, and thus, fewer lives are threatened when fires occur. Malls and marketplaces, family workshops, industrial and storage facilities, construction sites, restaurants, places of entertainment and dormitories are on the high fatality side, indicating that those places tend to be associated with relatively high-fatality fires.

Table 10
Decomposition of inertia among the year and fatality for the first two principles.[a]

Object	Quality	Mass	Inertia	Axis I			Axis II		
				Coordinate	Squared correlation	Contribution	Coordinate	Squared correlation	Contribution
Year									
2002	981	167	135	164	971	279	−17	10	3
2003	296	154	3	6	19	0	24	277	6
2004	1000	143	126	147	725	191	91	275	78
2005	432	129	26	−80	402	51	−22	29	4
2006	994	113	327	−94	90	62	−298	904	669
2007	1000	101	301	−256	654	411	186	346	234
2008	75	107	2	−1	0	0	26	75	5
2009	471	86	2	−29	413	5	11	58	1
Fatality									
3	873	535	113	18	41	11	−82	833	241
4–5	981	294	289	107	338	208	147	643	426
6–9	729	117	126	−154	477	173	−112	253	99
⩾10	980	54	395	−424	722	608	254	258	234

[a] All values are multiplied by 1000, and decimal points are omitted.

Table 11
Contingency table for provinces and fatality.

Province	Fatality				Total
	3	4–5	6–9	⩾10	
Anhui	10	4	0	0	14
Beijing	5	2	0	2	9
Chongqing	4	5	0	0	9
Fujian	23	10	3	3	39
Gansu	7	1	1	0	9
Guangdong	49	45	17	9	120
Guangxi	18	5	2	1	26
Guizhou	29	8	1	0	38
Hainan	4	0	3	0	7
Hebei	12	8	3	2	25
Henan	24	10	5	1	40
Heilongjiang	13	5	1	3	22
Hubei	9	5	3	0	17
Hunan	23	17	1	1	42
Jilin	7	1	1	3	12
Jiangsu	14	7	6	0	27
Jiangxi	15	8	3	1	27
Liaoning	9	6	3	1	19
Neimenggu	4	6	1	0	11
Qinghai	3	1	0	0	4
Shandong	9	5	3	4	21
Shanxi	9	4	5	0	18
Shaanxi	9	3	1	0	13
Shanghai	6	5	0	0	11
Sichuan	17	2	4	1	24
Tianjin	2	4	1	0	7
Xizang	3	1	0	0	4
Xinjiang	7	5	1	0	13
Yunnan	17	6	3	1	27
Zhejiang	32	27	14	7	80
Total	393	216	86	40	735

High occupant density and fire load may account for the fact that fires in malls and marketplaces tend to precipitate a high number of fatalities. Flammable finished materials, playing with fire, lack of evacuation facilities and high occupant density are four main reasons that lead to high-fatality fires in restaurants and places of entertainment (Lian, 2009; Zhang et al., 2006). In family workshops, to reduce the cost of production, many workers live, work and maintain storage at the same location. This alternative appears to be a predominant reason for many HCFs (Zhang, 2010; Liu, 2009). Fires that occur in industrial and storage facilities spread more widely than fires in other places (Holborn et al., 2002), which may explain the association between fires in industrial and storage facilities and high fatality rates. For construction sites, breaking safety regulations and a lack of firefighting equipment are two main factors that lead to large fires (Chen et al., 2009). For dormitories, the relatively high occupant density and a lack of firefighting equipment may be two main reasons for high-fatality fires. Hotels, welfare houses and hospitals are strongly associated with fatality level ⩾10. Because high-fatality fires in hotels often occur when people are asleep (Robert and Chan, 2000), sleep may be the underlying cause for high fatality. The relatively high vulnerability and low physical capability of occupants in welfare houses and hospitals may be a dominant reason for high-fatality fires. The place point 11 is close to the origin, which indicates that the fatality profile of place 11 is close to average profiles. Because "Others" contain various places, and the diversity of the category conceals the fatality characteristics of a given place, its fatality profile is similar to average profiles.

Thus, these qualitative or semi-quantitative results can provide useful instructions for fire prevention strategies. For example, given the tendency of high fatality rates, close attention should be paid to hotels, welfare houses and hospitals. Compared to other methods, the CA map is concise, vivid, easily understandable and especially suitable for fire officials to use.

3.2. Cause vs. fatality

The contingency table of fire causes and fatalities is shown in Table 3. The eigenvalues for the first two-dimensional axes are 0.139 and 0.119, with cumulative proportions of inertia equaling 51% and 88%, respectively.

The two-dimensional CA map of cause and fatality is shown in Fig. 2. We observe that fatality levels 3, 4–5, 6–9 and ⩾10 are relatively far from each other, which indicates that their cause profiles are quite different. Improper use of fire in daily life is a cause that has a point on the map relatively close to the point on the map for unknown cause, and thus, improper use of fire in daily life and unknown cause have similar fatality profiles. Similarly, electricity and arson are causes that have similar fatality profiles as well.

Table 4 provides the contributions of both fatality and cause variables for the first two dimensions. We can observe that fatality level ⩾10 is the primary contributor to axis I and that fatality level 3 is a secondary contributor. These two fatality levels are extreme in terms of their location on axis I. Between these two fatality levels, 75.6% of axis I is accounted for. As a result, axis I separates high fatalities on the right from low fatalities on the left. We can label axis I as the "fatality" axis. Then, the association between cause and fatality is obtained. Electricity, work-related tasks, arson, smoking, and other causes tend to precipitate fires with high fatality levels, and improper use of fires in daily life, playing with fire and unknown causes tend to cause fire with low fatality levels. Fires due to the work-related tasks should merit attention in fire prevention because they tend to precipitate high fatality levels. A typical example is the Shanghai high-rise fire, with 59 fatalities caused by electric welding. Compared to fires from other causes, fires caused by work-related tasks are easy to prevent by effective work safety regulations and laws. Given its high frequency and association with relatively high fatality rates, fires due to electrical causes should merit attention as well.

3.3. Time of day vs. fatality

The contingency table for the time of day and fatality is shown in Table 5. There are only 712 fire events available that contain information on the exact time of the fires. The eigenvalues for the first two-dimensional axes are 0.1634 and 0.063, with cumulative proportions of inertia equaling 83% and 95%, respectively.Fig. 3 shows the relative proximities of both fatality and time of day. Fatality levels 3, 4–5 and 6–9 form a group, indicating that their time profiles are similar. Fatality level ⩾10 is isolated and distinct. Times 8:00–11:59, 12:00–15:59 and 16:00–19:59 form one group, namely, the "day" group. Times 20:00–23:59, 0:00–3:59 and 4:00–7:59 form another group, namely, the "night" group. The times in each group have similar fatality profiles.

The contributions in Table 6 indicate that axis I is defined by fatality levels 4–5 and ⩾10. These two fatality levels are extreme in terms of their location on axis I. Then, axis I can be defined as the "fatality" axis and separates high fatalities on the left from low fatalities on the right. It is clear that the "day" group tends to be associated with high fatality rates and that the "night" group tends to be associated with low fatality rates. The reason for this finding is that people tend to gather in the daytime for various reasons; thus, a fire during the day would place more people in danger than one that occurs at night. By referring to Table 5, the frequency

Table 12
Decomposition of inertia among provinces and fatality for the first two principles.[a]

Object	Quality	Mass	Inertia	Axis I			Axis II		
				Coordinate	Squared correlation	Contribution	Coordinate	Squared correlation	Contribution
Province									
Anhui	815	19	23	431	801	58	−57	14	1
Beijing	690	12	35	−386	228	30	−549	462	75
Chongqing	555	12	18	163	63	5	454	492	51
Fujian	720	53	8	19	11	0	−154	709	25
Gansu	842	12	18	408	596	33	−262	246	17
Guangdong	982	163	67	−206	649	113	147	332	71
Guangxi	999	35	23	235	553	32	−211	446	32
Guizhou	974	52	80	459	857	178	−170	117	30
Hainan	12	10	1	−45	2	0	−112	11	2
Hebei	883	34	4	−133	878	10	10	5	0
Henan	707	54	8	146	687	19	−25	20	1
Heilongjiang	723	30	26	−131	92	8	−345	631	72
Hubei	441	23	6	110	142	5	159	298	12
Hunan	414	57	20	176	232	29	156	181	28
Jilin	968	16	88	−505	294	68	−765	674	193
Jiangsu	163	37	6	63	26	2	145	137	16
Jiangxi	993	37	2	77	884	4	27	109	1
Liaoning	691	26	3	−87	331	3	91	360	4
Neimenggu	877	15	28	12	0	0	538	877	87
Qinghai	881	5	8	467	822	19	−125	59	2
Shandong	987	29	68	−529	742	131	−304	245	53
Shanxi	64	24	3	2	0	0	138	64	9
Shaanxi	994	18	14	349	952	35	−74	43	2
Shanghai	533	15	13	263	267	17	263	266	21
Sichuan	574	33	28	172	128	16	−321	446	68
Tianjin	979	10	26	−106	26	2	646	953	80
Xizang	881	5	8	467	822	19	−125	59	2
Xinjiang	893	18	10	196	401	11	218	492	17
Yunnan	884	37	8	151	564	14	−113	320	10
Zhejiang	957	109	60	−277	860	137	93	97	19
Fatality									
3	999	535	168	177	637	273	−133	362	191
4–5	752	294	166	−69	41	23	289	711	494
6–9	230	117	62	−261	190	131	119	40	34
⩾10	922	54	314	−802	660	573	−506	263	281

[a] All values are multiplied by 1000 and decimal points are omitted.

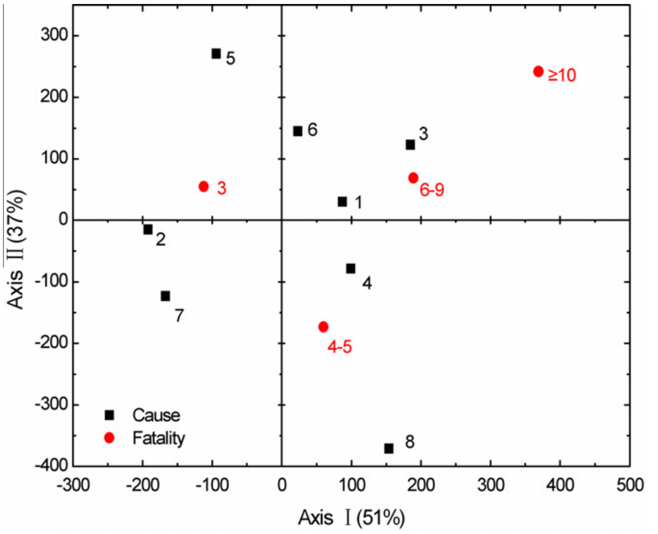

Fig. 2. CA map of cause vs. fatality. Cause: 1. Electricity, 2. Improperly using fire in daily life, 3. Work-related tasks, 4. Arson, 5. Playing with fire, 6. Smoking, 7. Unknown, 8. Others.

of fires occurring at night is considerably higher than that during the day, which indicates that the prevention strategies of night-time fires also merit consideration.

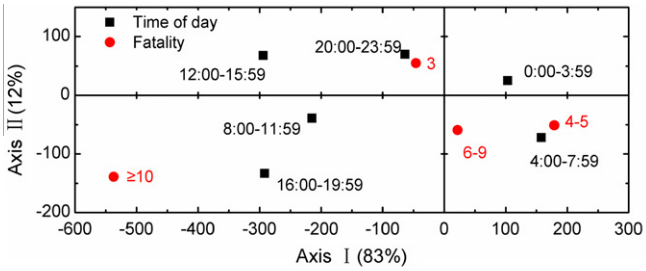

Fig. 3. CA map for the time of day vs. fatality.

3.4. Month vs. fatality

The contingency table for the month of the fire and fatalities is shown in Table 7. The eigenvalues for the first two-dimensional axes are 0.151 and 0.100, with cumulative proportions of inertia equaling 56% and 81%, respectively.

The display in Fig. 4 reveals that, by their proximities, some months have similar fatality profiles. March, November, December and October have similar profiles; September and February have similar profiles; May, June and January have similar profiles; and April and July have similar profiles. This similarity of fatality profiles might be explained by a similarity in climate at these times. The fatality points are spread out, indicating that their month profiles are quite different.

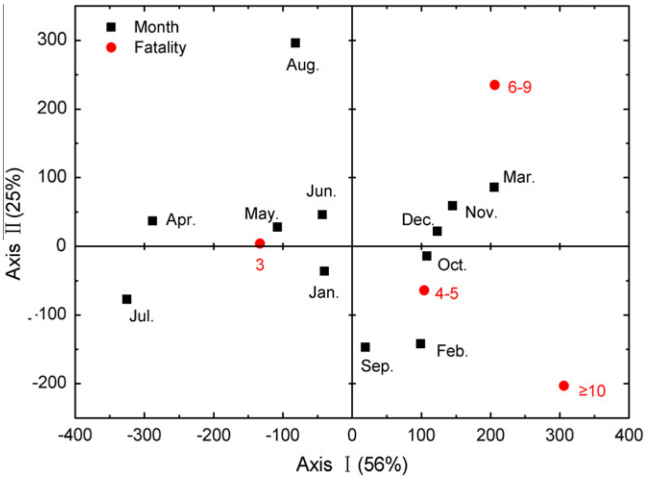

Fig. 4. CA map for the month vs. fatality.

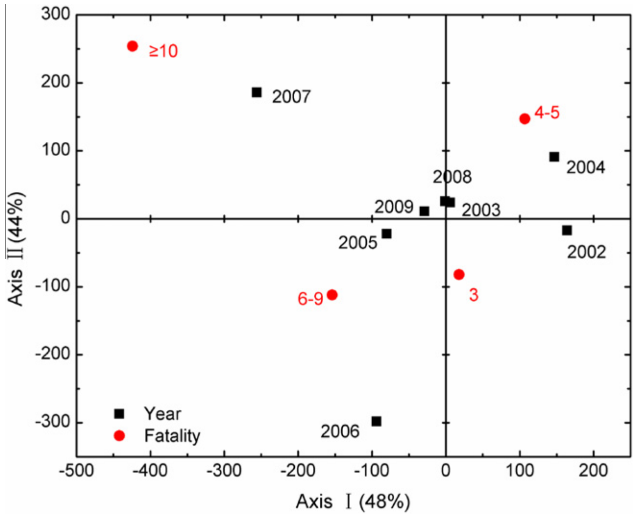

Fig. 5. CA map for the year vs. fatality.

According to Table 8, fatality levels 3 and ⩾10 are the primary and secondary contributors to axis I, respectively. Meanwhile, fatality level 3 is the lowest and fatality level ⩾10 is the highest value in axis I. As a result, we can label axis I as the "fatality" axis, and axis I separates low fatalities on the left from high fatalities on the right. Two categories of months are observed on the high fatality side of axis I: those that tend to be associated with fatality levels 4–5 and 6–9 (October, December, November, and March) and those that tend to be associated with fatality levels 4–5 and ⩾10 (September and February). As expected, most high-fatality months are in the cold season, such as the winter or the beginning of spring. On the left side of the space, we can identify months that tend to be associated with fatality level 3 (April, July, May, June and January), and August, which tends to be associated with fatality levels 3 and 6–9.

The association between the months and high fatality levels is of great importance for fire prevention and merits consideration. From Table 7 and Fig. 4, it is clear that fires with fatality level ⩾10 tend to be more strongly associated with February than the other months. The reason is that the Spring Festival, the most important festival in China, is usually in February. People like to ignite firecrackers during the Spring Festival, which is likely to cause fires (Yang et al., 2002).

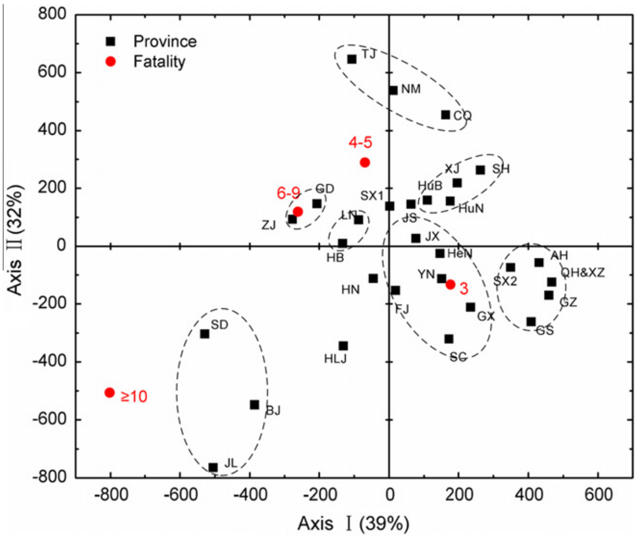

Fig. 6. CA map for the province vs. fatality. Province: AH Anhui, BJ Beijing, CQ Chongqing, FJ Fujian, GS Gansu, GD Guangdong, GX Guangxi, GZ Guizhou, HN Hainan, HB Hebei, HeN Henan, HLJ Heilongjiang, HuB Hubei, HuN Hunan, JL Jilin, JS Jiangsu, JX Jiangxi, LN Liaoning, NM Neimenggu, QH Qinghai, SD Shandong, SX1 Shanxi, SX2 Shaanxi, SH Shanghai, SC Sichuan, TJ Tianjin, XZ Xizang, XJ Xinjiang, YN Yunnan, ZJ Zhejiang.

Table 13
Fatality characteristics of the provinces.

Provinces	Fatality characteristics
Anhui, Gansu, Guizhou, Qinghai, Shaanxi, Xizang	Tendency for fires with fatality 3. A negligible amount of fires with fatality 6–9. No fire with fatality ⩾10
Fujian, Guangxi, Henan, Sichuan, Yunnan	Tendency for fires with fatality 3. A small amount of fires with fatality ⩾10
Hubei, Hunan, Shanghai, Xinjiang	Tendency for fires with fatality 4–5. A negligible amount of fires with fatality ⩾10
Chongqing, Neimenggu, Tianjin	Strong tendency for fires with fatality 4–5. No fires with fatality ⩾10
Hebei, Liaoning	Tendency for fires with fatality 4–5.
Hainan, Jiangsu, Shanxi	Strong tendency for fires with fatality 6–9. No fires with fatality ⩾10
Heilongjiang	Tendency for fires with both fatality 3 and ⩾10
Guangdong, Zhejiang	Tendency for fires with both fatality 4–5 and ⩾10
Beijing, Jilin, Shandong	Strong tendency for fires with fatality ⩾10
Jiangxi	Tendency for fires with average fatality profile

3.5. Year vs. fatality

The contingency table for the year of the fire and the fatality level is shown in Table 9. The eigenvalues for the first two-dimensional axes are 0.127 and 0.122, with cumulative proportions of inertia equaling 48% and 92%, respectively.

The CA map of the year-fatality data is shown in Fig. 5. The years 2003, 2005, 2008 and 2009 have similar fatality profiles that are close to the average fatality profile; 2002 and 2004 have similar fatality profiles. The fatality points are far apart, indicating that their year profiles are quite different. The contributions in Table 10 indicate that axis I is defined by fatality levels 4–5, 6–9 and ⩾10. Meanwhile, fatality levels 4–5 and ⩾10 are extreme in terms of their location on axis I. We define axis I as a "fatality" axis that separates high fatalities on the left from low fatalities on the right. Years 2002 and 2004 tend to be associated with low fatalities; years 2003, 2005, 2008 and 2009 tend to be associated with relatively high fatalities, compared to 2002 and 2004; and 2007 tends to be associated with very high fatalities.

We can discuss Fig. 5. from a dynamic perspective. It can be observed that fires in 2002 tended to cause low levels of fatalities. In 2003, fires tended to cause higher fatality rates compared to 2002. However, the fatality rate tendency in 2004 trended toward low fatality levels and was the same as in 2002. In 2007, fires more easily precipitated high fatality rates than in previous years. In 2009, fires tended to have a mid-level fatality rate, which falls between fatality tendencies in 2002 and 2007. Fig. 5 dynamically portrays fatality rate changes from 2002 to 2009. This result may provide insight into future work on fire prevention. For example, according to Table 9, the total number of HCFs steadily decreases. However, it is clear that HCFs in 2007–2009 tend to have higher fatality rates than those in 2002–2004. This observation indicates that fire prevention work should focus on high-fatality fires in future years.

3.6. Province vs. fatality

The contingency table for province and fatality is shown in Table 11. Three municipalities, Beijing, Tianjin and Shanghai, are considered in this work. The eigenvalues for the first two-dimensional axes are 0.221 and 0.209, with cumulative proportions of inertia equaling 39% and 71%, respectively.

The CA map of province-fatality data is shown in Fig. 6. It is apparent that province points have similar fatality profiles, which are highlighted by dashed circles. Provinces with similar fatality profiles have similar socioeconomic factors. For example, provinces Anhui, Qinghai, Xizang, Gansu, Guizhou and Shaanxi feature agricultural economies and undeveloped provinces, and their fatality profiles are very similar. Socioeconomic factors can greatly influence fire situations (Yang et al., 2005). Thus, similar socioeconomic factors may lead to similar fatality profiles. As shown in Table 12, the “Quality” of data for Shanxi, Jiangsu and Hainan province is very low, and thus, we cannot analyze the points of Shanxi, Jiangsu and Hainan province merely based on the CA map. Meanwhile, the fatality points 4–5 and 6–9 have similar province profiles and are different from fatality points 3 or ⩾10.

To analyze the associations between provinces and fatality, the character of the “fatality” dimension is defined. By referring to Table 12, fatality point ⩾10 is the primary contributor to axis I, with the lowest abscissa value, and fatality 3 is the secondary contributor with the highest abscissa value. As a result, axis I is defined as the “fatality” axis, which separates high fatalities on the left from low fatalities on the right. Many provinces are associated with low fatality levels, and the provinces of Guangdong, Zhejiang, Liaoning, Hebei and Heilongjiang tend to be associated with relatively high fatality levels; the provinces of Shandong, Beijing and Jilin tend to be associated with fatality level ⩾10. From a comprehensive analysis of Fig. 6. and Table 11, we can conclude the approximate characteristics of fire fatality situations in each province, the results of which are shown in Table 13. The fatality characteristics of HCFs by province provide useful information for the Fire Service Bureau and the Ministry of Public Security to characterize or classify the fire conditions of provinces and municipalities.

4. Comparisons between China and the United States

4.1. Place vs. fatality

For the sake of comparison, the contingency tables of two countries are combined, as shown in Table 14. Because the contingency table for place and fatality in the United States is a sparse matrix, CA cannot perform the analysis. Fortunately, the associations of fatality and place in the two countries are significantly different; thus, we are able to obtain results by making a direct comparison of the two contingency tables. In the following analysis, we focus primarily on the difference between the two countries. As shown in Table 14, HCFs with fatality level ⩾5 are concentrated on dwelling fires in the United States and are seldom in public areas. In contrast, in China, the most HCFs occurred in public areas and spread in many places. The difference between the two countries indicates the fire protection strategies that are adapted by specific areas as well as that the two countries have high complementarity and that a combined strategy may effectively control HCFs for the two countries. Specifically, because the fires in China with fatality level ⩾10 mainly occurred in public areas, it is worthwhile for the Chinese fire administration to study the American fire protection strategies for public areas.

Table 14
Contingency table for place and fatality in China and the United States.

Place	Fatality			
	5	6–9	⩾10	Sum
China				
Dwelling	18	13	3	34
Shop	11	10	2	23
Mall and marketplace	4	6	3	13
Family workshop	0	9	4	13
Industrial and storage	6	9	7	22
Construction site	1	2	1	4
Welfare house and hospital	1	6	1	8
Restaurant and entertainment	4	5	7	16
Hotel	1	3	4	8
Dormitory	0	1	1	2
Others	4	11	4	19
The United States				
Dwelling	58	40	5	103
Shop	0	0	0	0
Mall and marketplace	0	1	0	1
Family workshop	0	0	0	0
Industrial and storage	2	3	2	7
Construction site	0	0	0	0
Welfare house and hospital	0	0	3	3
Restaurant and entertainment	0	0	1	1
Hotel	1	0	1	2
Dormitory	0	0	0	0
Others	10	2	2	14
Sum	121	121	51	293

Table 15
Contingency table for the time of day and fatality in China and the United States.

Time of day	Fatality			
	5	6–9	⩾10	Sum
China				
0:00–3:59	19	30	8	57
4:00–7:59	12	16	5	33
8:00–11:59	5	6	5	16
12:00–15:59	4	5	5	14
16:00–19:59	2	6	7	15
20:00–23:59	8	12	7	27
The United States				
0:00–3:59	24	13	5	42
4:00–7:59	25	17	3	45
8:00–11:59	8	1	0	9
12:00–15:59	11	6	1	18
16:00–19:59	2	4	1	7
20:00–23:59	1	5	4	10
Sum	121	121	51	293

4.2. Time of day vs. fatality

Table 15 and Table 16 give respectively the contingency table and the decomposition of inertia for the time and fatality in China and the United States, and the corresponding CA map is shown in

Table 16
Decomposition of inertia among the time of day and fatality for the first two principles in China and the United States[a].

Object	Quality	Mass	Inertia	Axis I			Axis II		
				Coordinate	Squared correlation	Contribution	Coordinate	Squared correlation	Contribution
China									
0:00–3:59	1000	195	61	65	80	6	−221	920	328
4:00–7:59	1000	113	14	38	68	1	−141	932	77
8:00–11:59	1000	55	45	326	772	42	177	228	59
12:00–15:59	1000	48	68	424	751	61	244	249	99
16:00–19:59	1000	51	207	784	901	225	260	99	120
20:00–23:59	1000	92	42	277	997	51	16	3	1
The United States									
0:00–3:59	1000	143	88	−295	833	89	132	167	86
4:00–7:59	1000	154	108	−343	994	129	−27	6	4
8:00–11:59	1000	31	172	−896	852	176	373	148	148
12:00–15:59	1000	61	70	−435	992	83	40	8	3
16:00–19:59	1000	24	15	135	172	3	−296	828	72
20:00–23:59	1000	34	111	740	998	134	35	2	1
Fatality									
5	1000	413	400	−393	944	456	96	56	131
6–9	1000	413	132	127	299	48	−194	701	540
⩾10	1000	174	468	632	879	496	234	121	330

[a] All values are multiplied by 1000 and decimal points are omitted.

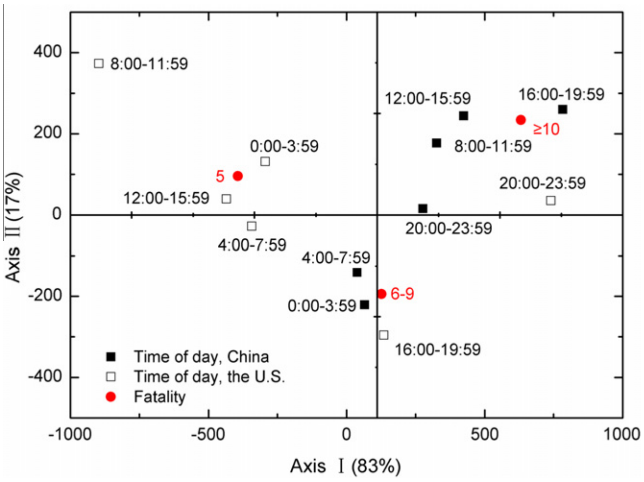

Fig. 7. CA map for the time of day vs. fatality in China and the United States.

Fig. 7. The eigenvalues for the first two-dimensional axes are 0.374 and 0.170, with cumulative proportions of inertia equaling 83% and 100%, respectively; these observations indicate that the first two dimensions can completely explain the total inertia.

By reference to Table 16, the axis I can be defined as the “fatality” axis and separates high fatalities on the right from low fatalities on the left. In general, it is clear that “US” points tend to be associated with a relatively low fatality level but that the “China” points tend to be associated with a relatively high fatality level. For the same time interval, the fatality profiles of the two countries present different features. For example, in the United States, 8:00–11:59 tends to be associated with fatality = 5; in contrast, this time interval tends to be associated with fatality level ⩾10 in China. The relative positions of time points within one country’s point set are to some degree different as well. For example, among American time points, 20:00–23:59 tends to be associated with the highest fatality level; however, among Chinese time points, 20:00–23:59 tends to be associated with moderate fatality levels. The reason for the difference in relative positions may be caused by the difference in the places where the HCFs occur.

Table 17
Contingency table for the month and fatality in China and the United States.

Month	Fatality			
	5	6–9	⩾10	Sum
China				
January	10	9	3	22
February	9	6	7	22
March	4	10	2	16
April	2	3	1	6
May	4	6	3	13
June	0	4	2	6
July	3	2	1	6
August	2	7	1	10
September	3	3	4	10
October	6	7	4	17
November	1	7	4	12
December	6	11	5	22
The United States				
January	11	7	0	18
February	7	4	4	15
March	8	5	3	16
April	9	2	1	12
May	5	3	0	8
June	3	2	0	5
July	6	5	0	11
August	3	3	1	7
September	5	4	3	12
October	5	6	1	12
November	3	1	1	5
December	6	4	0	10
Sum	121	121	51	293

4.3. Month vs. fatality

The month and fatality contingency table and the decomposition of inertia for the two countries are shown in Table 17 and Table 18, respectively. The contrasting CA results are shown in Fig. 8. The eigenvalues for the first two-dimensional axes are 0.334 and 0.224, with cumulative proportions of inertia equaling 66% and 100%.

According to Table 18, we can label axis I as the “fatality” axis, and axis I separates low fatalities on the right from high fatalities on the left. Fig. 8 shows that “China” points tend toward relatively high fatalities and that “the United States” points tend toward

Table 18
Decomposition of inertia among the month and fatality for the first two principles in China and the United States[a].

Object	Quality	Mass	Inertia	Axis I			Axis II		
				Coordinate	Squared correlation	Contribution	Coordinate	Squared correlation	Contribution
China									
January	1000	75	5	106	907	7	34	93	1
February	1000	75	72	−166	166	18	−373	834	176
March	1000	55	58	−197	207	18	385	793	136
April	1000	20	4	−114	385	2	145	615	7
May	1000	44	13	−225	995	20	16	5	0
June	1000	20	84	−814	926	118	230	74	18
July	1000	20	4	142	596	4	−117	404	5
August	1000	34	67	−246	177	18	530	823	161
September	1000	34	69	−425	510	54	−417	490	100
October	1000	58	10	−161	852	13	−67	148	4
November	1000	41	112	−685	980	167	99	20	7
December	1000	75	35	−275	922	49	80	78	8
The United States									
January	1000	61	95	499	920	133	147	80	22
February	1000	51	32	−20	4	0	−328	996	93
March	1000	55	14	119	324	7	−172	676	27
April	1000	41	110	620	820	137	−291	180	58
May	1000	27	45	520	945	64	125	55	7
June	1000	17	25	482	896	34	164	104	8
July	1000	38	47	398	717	52	250	283	39
August	1000	24	1	59	510	1	58	490	1
September	1000	41	11	−79	128	2	−206	872	29
October	1000	41	15	107	173	4	233	827	37
November	1000	17	19	259	339	10	−362	661	38
December	1000	34	51	482	896	69	164	104	15
Fatality									
5	1000	413	359	373	916	498	−113	84	89
6–9	1000	413	226	−150	237	81	270	763	506
⩾10	1000	174	416	−528	668	421	−372	332	405

[a] All values are multiplied by 1000 and decimal points are omitted.

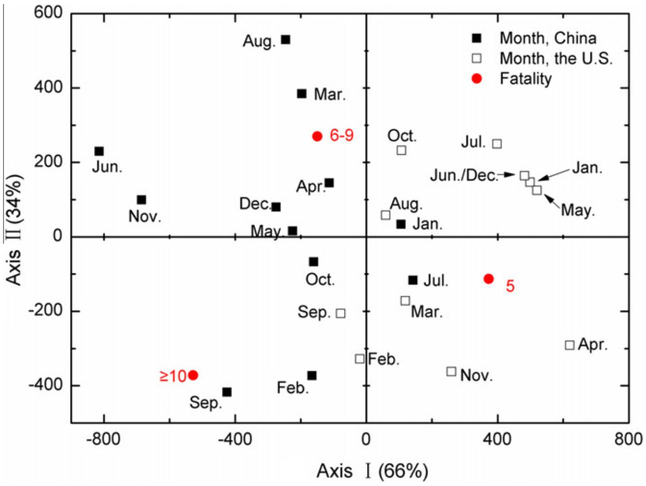

Fig. 8. CA map for the month vs. fatality in China and the United States.

Table 19
Contingency table for the year and fatality in China and the United States.

Year	Fatality			
	5	6–9	⩾10	Sum
China				
2003	7	14	6	27
2004	12	9	4	25
2005	6	10	7	23
2006	2	14	3	19
2007	5	10	9	24
2008	13	11	4	28
2009	5	7	4	16
The United States				
2003	9	12	3	24
2004	16	5	1	22
2005	6	6	3	15
2006	14	6	3	23
2007	10	8	2	20
2008	7	5	2	14
2009	9	4	0	13
Sum	121	121	51	293

relative low fatalities, which is similar to the results for the time of day. However, the relative positions of some month's points within one country's point set are similar between the two countries. For example, in both China and the United States, January and July tend to be associated with relatively low fatalities and are on the left side of axis I; September tends to be associated with relatively high fatalities, and its points are on the right side of axis I. A similarity exists between China and the United States climate, which might partially account for the relative similarity of the month points' relative positions within their country. The absolute similarity of the month points between the two countries is scarce. Only two February points have similar fatality profiles, and other months do not exhibit an obvious similarity.

4.4. Year vs. fatality

The year and fatality contingency table for the two countries is shown in Table 19. Table 20 gives the decomposition of inertia for the year and the fatality in the two countries. Fig. 9 shows the results of the comparison. The eigenvalues for the first two-dimensional axes are 0.356 and 0.170, with cumulative proportions of inertia equaling 82% and 100%, respectively.

Table 20
Decomposition of inertia among the year and fatality for the first two principles in China and the United States[a].

Object	Quality	Mass	Inertia	Axis I			Axis II		
				Coordinate	Squared correlation	Contribution	Coordinate	Squared correlation	Contribution
China									
2003	1000	92	58	307	966	69	58	34	11
2004	1000	85	10	−130	894	11	−45	106	6
2005	1000	78	78	353	805	77	−174	195	82
2006	1000	65	203	547	617	153	431	383	419
2007	1000	82	176	490	721	156	−305	279	266
2008	1000	96	8	−111	959	9	23	41	2
2009	1000	55	21	227	874	22	−86	126	14
The United States									
2003	1000	82	19	39	44	1	185	956	97
2004	1000	75	202	−645	997	247	−34	3	3
2005	1000	51	2	39	324	1	−56	676	6
2006	1000	78	81	−380	905	90	−123	95	42
2007	1000	68	22	−202	809	22	98	191	23
2008	1000	48	10	−176	985	12	−22	15	1
2009	1000	44	111	−609	952	130	137	48	29
Fatality									
5	1000	413	462	−415	990	561	−42	10	26
6–9	1000	413	216	229	644	171	170	356	416
⩾10	1000	174	321	441	679	268	−303	321	558

[a] All values are multiplied by 1000 and decimal points are omitted.

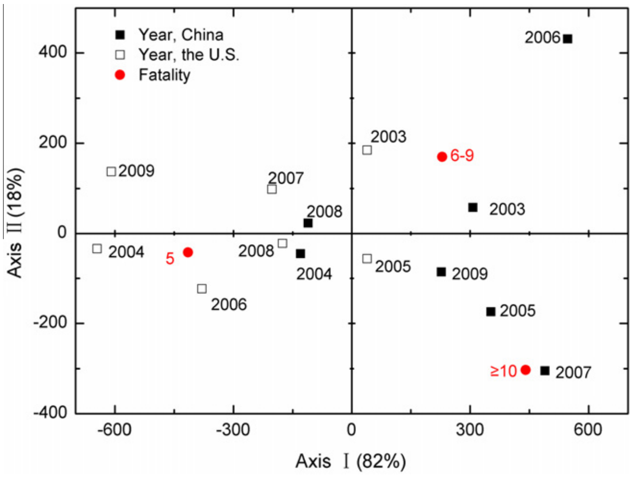

Fig. 9. CA map for the year vs. fatality in China and the United States.

By reference to Table 20, we can label axis I as the "fatality" axis, which separates high fatalities on the right from low fatalities on the left. In contrast, "China" points generally tend to be associated with higher fatality than "the United States" points. Among the two countries' point sets, there is little similarity. There are only two 2004 points that have similar relative positions within their country points and two 2008 points that have close fatality profiles. From a dynamic perspective, the total number of HCFs and the percent of fires with relatively high fatalities decrease with time in the United States. In contrast, in China, the total number of HCFs fluctuates, and HCFs still tend to be associated with relatively high fatalities. Therefore, we conclude that both the HCF situations and the trends in the United States are better than those in China.

4.5. Discussion

According to the above analysis, the most notable difference between the two countries is that HCFs in China tend to be associated with higher fatality levels than HCFs in the United States. Public structures generally have higher occupant density and more occupants than residential structures, which implies that public structures have higher probability to experience HCF when fire occurs. As shown in Table 14, more than half of the Chinese HCFs (66%) occur in public structures; by contrast, the majority of the American HCFs (78%) occur in residential structures. Therefore, the difference in the places where the fires occur could account partially for the different fatality tendency of HCFs in these two countries. The difference may be attributed to a poor fire safety management situation of public structures in China, and to a generally higher fire risk in residential structures in the United States.

In China, many public structures, especially for entertainmental places, frequently infringe the requirements of the fire safety regulations. For instance, in order to prevent thievery, the managers of those public structures probably lock the evacuation exits that are not in daily use; neither the smoke alarms nor the automatic suppression equipments are working due to the lack of maintenance. In comparison, these situations are infrequent in American public structures. In the United States, not only the fire departments but also the insurance companies play a very important role in supervising the fire safety situations of public structures. Without insurance companies involved in supervising and guiding daily fire safety works, Chinese fire departments have to difficultly fulfill all fire-related tasks, which may enlarge the gap between two countries.

By comparison with China, the American residential structures have higher fire risk, which may be due, in part, to the following factors: unprotected wood-frame constructions, children and residential structure density. In the United States, many residential structures are wood-frame constructions and most catastrophic multiple-death residential fires occur in unprotected wood-frame constructions (Badger, 2004, 2005, 2006, 2007, 2008, 2009, 2010). An unprotected wood-frame house is more conducive to fire spread than a house built with incombustible materials. Conversely, most Chinese residential structures are built with incombustible materials such as concrete and brick, which substantially decreases the fire risk. Given the fact that children under 6 years old account for a large part of the fire facilities in American HCFs, the factor of children may influence the gap of the fire risk of residential structures between the two countries. In China, most families only have one child because of the Family Planning Policy, but many American families have several children. Because the cognitive abilities of child are limited, the risk of death and injury from fire rises (US Fire Administration National Fire Data Center, 2011). As a result, more children make the residential structures in the United States have higher fire risk than that in China. Previous study (Hall, 2003) indicates that the United States has a higher incidence of single parent families and a higher incidence of lack of child supervision than Japan. Because China and Japan have some similarity in family tradition, Chinese children may be better supervised than American children as well. Possibly this is part of the reason that the factor of children increases the fire risk of residential structure in the United States. The density of residential structures also has an effect. In China, the density of residential structures is considerably higher than that in the United States and a house may have more neighbors. When fire occurs, adjacent neighbors would quickly recognize the fire and become volunteer firefighters who often play a decisive role in rescuing fatalities and controlling fire. Therefore, the relative lack of neighbors' help can increase the fire risk of residential structures in the United States.

5. Conclusions

Analyzing past fire accidents enables us to understand the ways in which the accidents occur and provides useful input for the development of preventive strategies. The analysis performed in the present study is aimed at investigating the associations between six influential factors and the fatality levels of HCFs from 2002 to 2009 in China. CA was used as the analysis technique to explore the associations. The main results of this study are shown in Table 21. For the sake of simplicity, in Table 21, the associations are portrayed only by relatively low and high fatality rates. The power of CA in exploring the relationships among categorical variables is demonstrated. The associations between influence factors and fatality have practical implications. There are some factors that tend to be associated with very high fatality levels; these factors have sufficient importance in fire prevention that they merit special attention. The results obtained in the comparisons of HCFs in China and the United States clearly demonstrate that Chinese HCFs tend to be associated with higher fatality levels than HCFs in the United States. For four influence factors, place, time of day, month and year, HCFs in China and the United States present quite different fatality profiles. The two most significant comparison results can be stated as follows: (1) HCFs in China and the United States are inclined to occur in public areas and dwellings, respectively; (2) In the United States, the trend of HCFs is improving, but HCFs in China have not been effectively controlled.

Table 21
Associations between influence factors and fatality.

Factor	Fatality	
	Relatively low	Relatively high
Place	Dwelling, Shop, Others	Mall and marketplace, Family workshop Industrial and storage, Construction site Welfare house and hospital, Restaurant and entertainment, Hotel Dormitory
Cause	Improperly using fire in daily life, Playing with fire, Unknown	Electricity, Work-related tasks, Arson, Smoking, Others
Time of day	20:00–23:59, 0:00–3:59, 4:00–7:59	8:00–11:59, 12:00–15:59, 16:00–19:59
Month	January, April, May, June, July, August	September, October, November, December, February, March
Year	2002, 2004	2003, 2005, 2006, 2007, 2008, 2009
Province	Anhui, Chongqing Fujian, Gansu, Guangxi, Guizhou, Henan, Hubei, Hunan, Jiangxi, Neimenggu, Qinghai, Shaanxi, Shanghai, Sichuan, Tianjin, Xinjiang, Xizang, Yunnan	Beijing, Guangdong, Hainan, Hebei, Heilongjiang, Jiangsu, Jilin, Liaoning, Shandong, Shanxi, Zhejiang

To control the HCFs, prevention measures will need to be carefully targeted at the influence factors that tend to be associated with high fatality levels or account for the majority of HCFs. The following recommendations may be useful for the HCFs controlling in China. Efforts must be made to control the HCFs in public structures, especially HCFs in hotels, welfare houses and hospitals. Properly maintaining smoke alarms and automatic suppression equipments, routinely inspecting the means of egress, and strictly supervising by the fire departments may efficiently prevent HCFs in public structures. As shown in Table 3, electricity and improperly using fire in daily life related fires account for 33% and 25% of the HCFs, as a result, these two causes merit special attention to reduce their incidence. Most electricity-related fires can be prevented by properly installing electrical circuit and appliances. Frequently national fire safety campaigns that address lifestyle strategies of safety using fires may significantly reduce the incidence of HCFs caused by improperly using fire in daily life.

It is our hope that this study will be useful to the further study, understanding, and control of HCFs. However, the lack of underreporting analysis is still a practical limitation. Further research will be devoted to removing this limitation.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (No. 91024027).

References

Amoros, E., Martin, J., Laumon, B., 2006. Under-reporting of road crash casualties in France. Accident Analysis and Prevention 38, 627–635.

Badger, S.G., 2010, Catastrophic multiple-death fires for 2009. National Fire Protection Association, Fire Analysis and Research Division. <http://www.nfpa.org/categoryList.asp?categoryID=413&URL=Research/Fire%20reports/Overall%20fire%20statistics> (05.08.11).

Badger, S.G., 2009, Catastrophic multiple-death fires for 2008. National Fire Protection Association, Fire Analysis and Research Division. <http://www.nfpa.org/categoryList.asp?categoryID=413&URL=Research/Fire%20reports/Overall%20fire%20statistics> (05.08.11).

Badger, S.G., 2008, Catastrophic multiple-death fires for 2007. National Fire Protection Association, Fire Analysis and Research Division. <http://www.nfpa.org/categoryList.asp?categoryID=413&URL=Research/Fire%20reports/Overall%20fire%20statistics> (05.08.11).

Badger, S.G., 2007, Catastrophic multiple-death fires for 2006. National Fire Protection Association, Fire Analysis and Research Division. <http://www.nfpa.org/categoryList.asp?categoryID=413&URL=Research/Fire%20reports/Overall%20fire%20statistics> (05.08.11).

Badger, S.G., 2006, Multiple-death fires for 2005. National Fire Protection Association, Fire Analysis and Research Division. <http://www.nfpa.org/categoryList.asp?categoryID=413&URL=Research/Fire%20reports/Overall%20fire%20statistics> (05.08.11).

Badger, S.G., 2005, Catastrophic multiple-death fires in the United States – 2004. National Fire Protection Association, Fire Analysis and Research Division. <http://www.nfpa.org/categoryList.asp?categoryID=413&URL=Research/Fire%20reports/Overall%20fire%20statistics> (05.08.11).

Badger, S.G., 2004, Catastrophic multiple-death fires in the United States – 2003. National Fire Protection Association, Fire Analysis and Research Division. <http://www.nfpa.org/categoryList.asp?categoryID=413&URL=Research/Fire%20reports/Overall%20fire%20statistics> (05.08.11).

Barillo, D.J., Goode, R., 1996. Fire fatality study: demographics of fire victims. Burns 22, 85–88.

Chen, J.R., Zhang, Y.T., Sun, J.L., 2009. Fire prevention and contingency rescue for project under construction. Journal of Wuhan Institute of Technology 31, 48–51.

Chen, Z.J., Wang, F.L., Lu, S.X., Fang, Z.B., 2007. Cluster analysis of fire data in China. Engineering Science 9, 86–94.

Clausen, S.E., 1998. Applied Correspondence Analysis: An Introduction. Sage University Papers Series on Quantitative Applications in the Social Science, 07-121. Sage Publications, Thousand Oaks, CA.

Corman, H., Ignall, E.J., Rider, K.L., Stevenson, A., 1976. Fire casualties and their relation to fire company response distance and demographic factors. Fire and Technology 12, 193–203.

Duncanson, M., Woodward, A., Reid, P., 2002. Socioeconomic deprivation and fatal unintentional domestic fire incidents in New Zealand 1993–1998. Fire Safety Journal 37, 165–179.

Elder, A.T., Squires, T., Busuttil, A., 1996. Fire fatalities in elderly people. Age and Ageing 25, 214–216.

Fire Service Bureau, 2003. Fire Statistical Year Book of China 2003. China Personnel Publishing House, Beijing, pp. 50–252.

Fire Service Bureau, 2005a. China Fire Services 2004. China Personnel Publishing House, Beijing, pp. 157–212.

Fire Service Bureau, 2005b. China Fire Services 2005. China Personnel Publishing House, Beijing, pp. 315–397.

Fire Service Bureau, 2006. China Fire Services 2006. China Personnel Publishing House, Beijing, pp. 277–355.

Fire Service Bureau, 2007. China Fire Services 2007. China Personnel Publishing House, Beijing, pp. 409–476.

Fire Service Bureau, 2008. China Fire Services 2008. China Personnel Publishing House, Beijing, pp. 489–517.

Fire Service Bureau, 2009. China Fire Services 2009. China Personnel Publishing House, Beijing, pp. 207–241.

Fire Service Bureau, 2010. China Fire Services 2010. China Personnel Publishing House, Beijing, pp. 249–274.

Garson, G.D., 2008. Correspondence Analysis, Statnotes: Topics in Multivariate Analysis. <http://www2.chass.ncsu.edu/garson/pa765/correspondence.htm> (March 2008).

Guo, T.N., Fu, Z.M., 2007. The fire situation and progress in fire safety science and technology in China. Fire Safety Journal 42, 171–182.

Hall, J.R., 2003, Fire in the US and Japan. Nation Fire Protection Association, Fire Analysis and Research Division. <http://www.nfpa.org/itemDetail.asp?categoryID=421&itemID=18252&URL=Research/Fire%20reports/International%20comparisons> (08.11.11).

Hasofer, A.M., Thomas, I., 2006. Analysis of fatalities and injuries in building fire statistics. Fire Safety Journal 41, 2–14.

Higgs, N.T., 1990. Practical and innovative uses of correspondence analysis. The Statistician 40, 183–194.

Hoffman, D.L., Franke, G.R., 1986. Correspondence analysis: graphical representation of categorical data in marketing research. Journal of Marketing Research 23, 213–217.

Holborn, P.G., Nolan, P.F., Golt, J., 2003. An analysis of fatal unintentional dwelling fires investigated by London Fire Brigade between 1996 and 2000. Fire Safety Journal 38, 1–42.

Holborn, P.G., Nolan, P.F., Golt, J., Townsend, N., 2002. Fires in workplace premise: risk data. Fire Safety Journal 37, 303–327.

Johnson, R.A., Wichern, D.W., 2007. Applied Multivariate Statistical Analysis, sixth ed. Pearson Prentice Hall, Upper Saddle River, New Jersey, pp. 716–726.

Jordan, L.B., Squires, T.J., Busuttil, A., 1999. Incidence trends in house fire fatalities in eastern Scotland. Journal of Clinical Forensic Medicine 6, 233–237.

Li, H.J., 2010. Statistic analysis of the serious fire disasters in China during 2000–2008. China Public Security. Academy Edition 18, 64–69.

Lian, C.H., 2009. Fire investigation of LATIN BAR in Changle and the analysis of factors contributing to the fire. Fire Science and Technology 28, 705–707.

Liu, Z.Q., 2009. Cause and Countermeasures to fire risk in 3-in-1 sites. Journal of Institute of Disaster Prevention Science and Technology 11, 44–47.

Park, M., Lee, J.W., Kim, C., 2007. Correspondence analysis approach for finding allele associations in population genetic study. Computational Statistics & Data Analysis 51, 3145–3155.

Pusha, S., Gudi, R., Noronha, S., 2009. Polar classification with correspondence analysis for fault isolation. Journal of Process Control 19, 656–663.

Robert, D., Chan, D.H., 2000. Fires in hotel rooms and scenario predictions. International Journal of Contemporary Hospitality Management 12, 37–45.

Sourial, N., Wolfson, C., et al., 2010. Correspondence analysis is a useful tool to uncover the relationships among categorical variables. Journal of Clinical Epidemiology 63, 638–646.

Squires, T., Busuttil, A., 1995. Child fatalities in Scottish house fires 1980–1990: a case of child neglect. Child Abuse & Neglect 19, 865–873.

US Fire Administration National Fire Data Center, 2011. Fire Risk in 2007. <http://www.usfa.fema.gov/statistics/reports/casualties.shtm> (08.11.11).

Yang, L.Z., Chen, H., Yang, Y., Fang, T.Y., 2005. The effect of socioeconomic factors on fire in China. Journal of Fire Sciences 23, 451–461.

Yang, L.Z. et al., 2002. Fire situation and fire characteristic analysis based on fire statistics of China. Fire Safety Journal 37, 785–802.

Zhang, G.J., 2010. Fuzzy synthetic evaluation for fire risk in a spot of three functions. Journal of Chinese People's Armed Police Force Academy 26, 30–32.

Zhang, S.P., Xing, Y.G., Wan, J., Bai, L., 2006. Countermeasures to and performance-based assessment of KTV fire hazards. Fire Science and Technology 25, 756–758.