ELSEVIER

Journal of Hazardous Materials 71 (2000) 239–252

Journal of Hazardous Materials

www.elsevier.nl/locate/jhazmat

Statistical analysis of domino chemical accidents

S.P. Kourniotis *, C.T. Kiranoudis, N.C. Markatos

Department of Chemical Engineering, National Technical University of Athens, Zografou Campus, Athens 157 80 GR, Greece

Abstract

A set of chemical accidents is retrieved from the literature and classified with regard to the substance involved and whether domino effects are present. This set of accidents and each of the classes defined are statistically analyzed with respect to its severity and comparison is made between domino and non-domino accidents. The analysis reveals that each accident category shows characteristic patterns in terms of fatalities caused and domino effects likelihood. Moreover, chemical accidents severity frequencies are described by using a two-parameter, revised form of the Pareto probability density function. The range within which the values of the parameters lie is investigated using Bayesian inference. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Domino chemical accidents; Statistical analysis; Bayesian inference

1. Introduction

Domino chemical accidents are defined as those situations where a chemical accident becomes the initiating event of one or more other accidents, thus, increasing the severity of the off-site consequences. Analysis of such cases is a topic of increasing interest due to the fact that they tend to increase the severity of the consequences, and in addition they are difficult to be predicted and effectively managed. Up to date, relevant contingency plans do not usually refer to domino effects. The European Council Directive 96/82/EC (Seveso II) [1] requires the identification of those establishments

Abbreviations: CEPPO: Chemical Emergency Preparedness and Prevention Office; HSE: Health and Safety Executive; MARS: Major Accident Reporting System; PDF: Probability Density Function

* Corresponding author. E-mail: kyr@orfeas.chemeng.ntua.gr

PII: S0304-3894(99)00081-3

where the likelihood and the possibility of consequences of a major accident may be increased due to domino effects.

Chemical accidents are combinations of causes and consequences. From a societal point of view, the severity of an accident is primarily defined by the magnitude of its consequences, that is, by the harm it can cause to health. On the other hand, the causes of an accident are of primary importance also, when it is analyzed from the engineering point of view, considering that different causes may produce different magnitudes of consequences. Chemical accidents resulting in fatalities, and dominos in particular, are rare events and information about them includes a lot of uncertainties. In spite of these limitations, probability theory is the only rational way that is available for handling uncertainty [2]. Although, in some cases, the results may be qualitatively obvious, it is important to have a mathematical model describing them, built on a number of assumptions and parameters whose numerical values can be obtained.

Chemical accidents are usually examined by constructing fatalities–frequency (f–N) curves of a selected set of accidents and investigating the patterns of those curves to infer the likelihood of accident severity. By using this method, the assumption is made that the severity of an accident depends mainly on the number of fatalities caused. Although the number of fatalities is representative of the severity up to a point, efficient emergency management depends on other factors too, accident duration and economic loss being two of them.

Domino accidents have been commonly regarded as more severe than the rest. In this study, the review of chemical accident case histories performed is used to investigate differences of domino and non-domino accidents. The number of fatalities is adopted as a probable indicator of such differences.

Furthermore, statistical analysis is used to predict the likelihood of chemical accidents for both domino and non-domino cases. It should be kept in mind that different sets of accidents may be used and have been used by researchers, in addition to the fact that accident behavior and emergency management change through time due to technological and societal changes, respectively. Accidents themselves are considered to be discrete events occupying both time and space. In this case, the limitations imposed due to the quality of available historical data should be taken into account [3]. Bayesian analysis is used to overcome the above limitations by predicting the range of values within which lie the parameters describing the fatalities frequency of chemical accidents.

2. Characteristics of the chemical accidents set

A set of 207 major chemical accidents is retrieved from relevant literature, competent authorities reports (CEPPO, HSE) and well established accident databases (MARS) [4–9]. A list of important accidents is extracted from the sample and is shown in Appendix A. The chronological and geographical distributions are regarded as significant aspects of the sample as improved technologies, working practices and emergency management procedures followed by different countries tend to change industrial accidents attitude through time and space. Narrowing the range of time within which the accidents of the sample lie, and limiting the countries to those which are considered to

Table 1
Geographical distribution of accidents

Geographic area	Number of accidents	%
W. Europe	127	61.4
N. America	57	27.5
Other	23	11.1
Total	207	100

be economically and/or industrially developed, are qualitative measures to ensure a minimum degree of homogeneity to the data of the sample.

All the accidents of the sample refer to economically developed countries, compared by the size of their gross national product. The geographical distribution of the sample is given in Table 1. All the accidents occurred during the last 40 years, while the majority of them occurred within the last 2 decades. The chronological distribution of the sample is given in Table 2.

The sample includes accidents that have occurred both in fixed installations (process and storage facilities) and during transportation. They include all methods of transportation (pipeline, rail, road and shipping) and a variety of installation types including oil refineries, petrochemicals, fertilizers manufacturing, fine chemicals industries, etc.

The major assumption is that the severity of an industrial accident depends on the chemical substance primarily involved in it. Thus, the accidents of the sample are classified into categories depending on the characteristics of the chemical substances primarily involved in them. These categories involve:

liquid fuels (e.g. crude oil, petrol, kerosene, naphtha);
vapor hydrocarbons (hydrocarbons with up to four atoms of carbon in their molecule);
toxic substances (e.g. chlorine, ammonia, pesticides);
miscellaneous (all substances not included in the above categories).

The categories defined above, although not equivalent, are considered to be representative of the data contained in the sample. The last category is a dummy one and has been defined so as to include all the accidents of the sample that could not be included in the other three categories. Moreover, the number of accidents that belong to each category is approximately assumed to be a percentage representation of the accidents occurring worldwide.

Table 2
Chronological distribution of accidents

Time period	Number of accidents	%
Up to 1969	17	8.2
1970 to 1979	42	20.3
1980 to 1989	74	35.7
1990 to 1998	74	35.7
Total	207	100

The probability of domino accident within the data of each category is estimated. More specifically, the number of accidents within each category, the number of accidents known to have caused at least one domino, and the number of accidents known to have caused more than one domino are given in Table 3. The term probability is used as a measure of belief, and it should not be interpreted as the limit of the relative frequency of occurrence [2].

As can be noted from Table 3, each category shows a different trend regarding the appearance of domino accidents. More specifically, it seems that substance properties and the way of its handling and/or storage greatly affect the probability to have domino accidents once an initiating accident has occurred. The probability of a domino accident for all the cases contained in the sample is approximately 0.39. Flammable substances (hydrocarbons) tend to cause domino accidents more often. Vapor hydrocarbons, usually stored and transferred under pressure, is the sample category with the larger likelihood to cause one or more dominos (0.58). On the other hand, toxic substances do not usually provoke domino effects (0.16). These results can be logically explained by the fact that violent phenomena like fires and explosions, usually caused by flammable and/or explosive substances, can easily cause mechanical failures in the vicinity of the initiating accident. On the other hand, toxic clouds, although they generally tend to occupy a wider geographical area, do not usually cause direct mechanical failures.

All the accidents within the sample, as well as those for each category, are statistically analyzed in order to determine specific patterns that each category follows in terms of severity. Measure of severity is considered to be the number of fatalities caused. The differences between domino and non-domino accidents within each category are also examined.

The f–N curves of the sample and each category are constructed and shown in Fig. 1. They are based on accidents involving fatalities ranging from 1 to 600. In this figure, the conditional probability of an accident (that occurred and resulted in fatal consequences) to result in N or more fatalities ($P[x \geq N | N \geq 1]$) is plotted against the number N of fatalities. The f–N curve of the sample, as well as of each category examined, has a slope of about -1, which is in agreement with the work presented at the second Canvey Report [10] and by Haastrup and Brockhoff [11]. This slope indicates that the probability of a chemical accident is reduced logarithmically in relation to its severity. However, differences seem to exist among accidents involving different substances. In the range of

Table 3
Likelihood of domino accidents

Category	Liquid fuels	Vapor hydrocarbons	Toxic substances	Miscellaneous substances	Total
number of accidents	43	50	45	69	207
number of accidents with at least one domino (probability)	21 (0.488)	29 (0.580)	7 (0.156)	23 (0.333)	80 (0.386)
number of accidents with at least two dominos (probability)	8 (0.186)	14 (0.280)	2 (0.044)	10 (0.145)	34 (0.164)

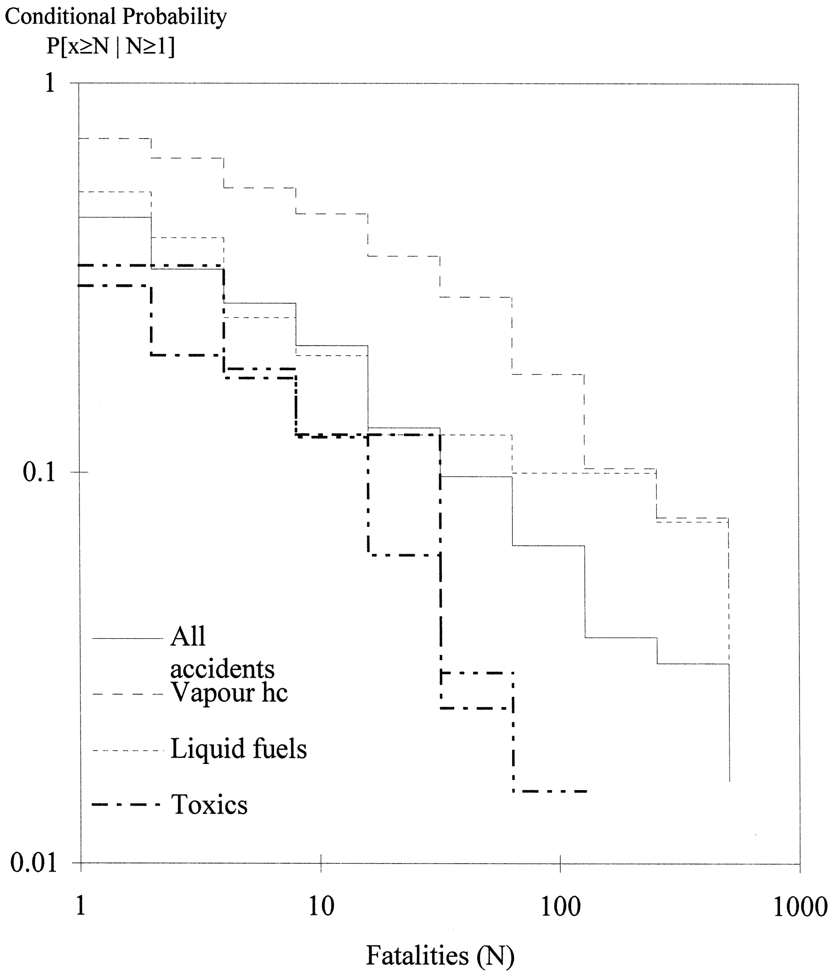

Fig. 1. $f-N$ curves of the accident categories.

up to approximately 250 fatalities, accidents involving vapor hydrocarbons are more probable (the corresponding curve is less steep), followed by those involving liquid fuels, which become more probable than all other categories for cases exceeding 250 fatalities. On the other hand, the dummy category of miscellaneous substances, generally, seems less probable to cause high-consequence accidents. The *f–N* curve of accidents involving toxic substances lies within the range of 1–100 fatalities and have a somewhat steepest slope.

f–N curves are constructed for the domino accidents of the set in Fig. 2. Comparison reveals that domino accidents of the set have curves of larger severity (higher probability to result to *N*, or more, fatalities). The same stands for the categories of liquid fuels and miscellaneous substances. Vapor hydrocarbons differ in the sense that the data of

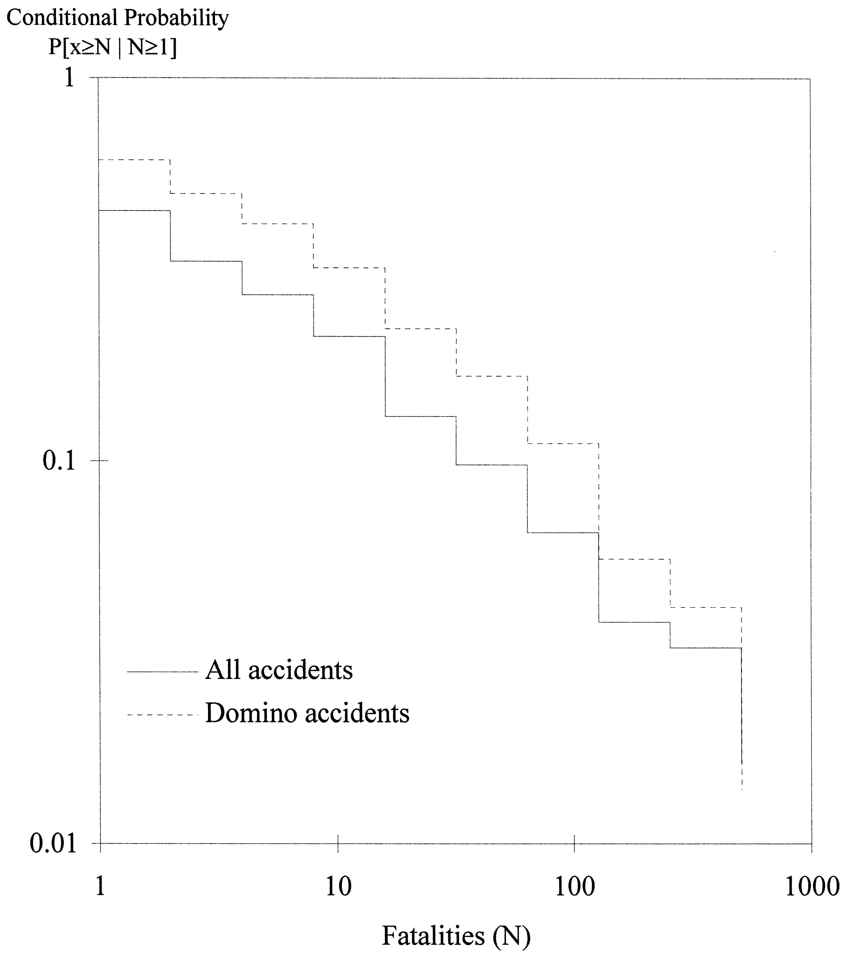

Fig. 2. $f-N$ curve comparing all domino accidents.

this category do not give a clear indication whether domino accidents are expected to be more severe. In the case of toxics, the size of the sample of domino cases does not permit to obtain conclusive results.

3. Mathematical modeling

The evaluation of a probability density function (PDF) that represents the data of natural as well as man-caused catastrophes becomes an increasing challenge during recent years. In most cases found in literature, the exponential, log-normal and Pareto distributions have been used for this purpose [12]. The exponential distribution has been used by insurance companies, to predict the frequency of low-probability, high-consequence catastrophes, in terms of fatalities, but tends to under-predict low-consequence incidents. On the other hand, the log-normal distribution has been the basis to derive probit functions in order to predict human injuries, given that the accident consequences are known in terms of physical quantities (e.g. thermal radiation flux) [4]. Log-normal distribution describes acceptably accidents resulting in few fatalities, but tends to over-predict the frequency of low-probability high-consequence accidents. Pareto PDF belongs to the family of the exponential shape distributions. The form of the distribution is [13]:

$$f(z|\alpha) = \frac{\alpha Z_0^{\alpha}}{z^{\alpha+1}}, \quad \text{for } z \geq Z_0 \tag{1}$$

where Z_0 = parameter equal to the minimum event size of interest, z = incident size (number of fatalities) and α = constant.

To take into account the complete distribution of sizes down to zero fatalities, a revised version of Pareto PDF is used by introducing a second parameter K:

$$f(z|\alpha,K) = \frac{\alpha K^{\alpha}}{(K+z)^{\alpha+1}}, \quad \text{for } z \geq 0 \tag{2}$$

where K = location parameter and α = scale parameter.

The corresponding cumulative distribution function is:

$$f(z) = 1 - \left(\frac{K}{K+z}\right)^{\alpha}. \tag{3}$$

The parameters α and K are estimated using the maximum likelihood (ML) estimation method [14]. The likelihood function, L, is defined as the PDF describing the density of the parameter α, if the parameter itself is conceived as a random variable taking values from a parameter space. The maximum likelihood estimator of α is defined as that value of α that maximizes L. In other words, the ML estimator of α is the value of α which maximizes the probability of the observed data. The maximum likelihood function of the Pareto PDF, for n observations $(z_1,\ldots,z_n)$ is:

$$L(z_1,z_2,\ldots,z_n|\alpha,K) = \prod_{i=1}^{n} f(z_i|\alpha,K) \tag{4a}$$

or

$$\ln L(z_1,z_2,\ldots,z_n|\alpha,K) = \sum_{i=1}^{n} \ln f(z_i|\alpha,K)$$
$$= n\ln\alpha + \alpha n \ln K - (\alpha+1)\sum_{i=1}^{n}\ln(K+z_i). \tag{4b}$$

The estimators of α and K are calculated by solving the system of equations:

$$\frac{\partial \ln L}{\partial \alpha} = 0 \tag{5}$$

and

$$\frac{\partial \ln L}{\partial K} = 0 \tag{6}$$

while the estimators must fulfill the conditions:

$$\left.\frac{\partial^2 \ln L}{\partial \alpha^2}\right|_{\alpha=\hat{\alpha}} < 0 \tag{7}$$

and

$$\left.\frac{\partial^2 \ln L}{\partial K^2}\right|_{K=\hat{K}} < 0 \tag{8}$$

respectively.

The calculations described above were applied to all accident categories and the validity of the estimated distributions are tested using the Kolmogorov–Smirnov test [13].

The calculated values of α and K for all the accidents of the sample are $\alpha = 0.79$ and $K = 3.54$. The Kolmogorov–Smirnov test showed that the data fit well to the PDF Pa (0.79, 3.54). The calculated estimators of α and K for all accident categories and domino cases are given in Table 4. The estimated theoretical distributions, for each category, are plotted in Fig. 3 for the range of 1–600 fatalities, where the sample data lie, and are extrapolated up to 1000 fatalities. The comparative severity of each category depends on the slope of the curve and, consequentially on both the parameters α and K of the distribution. The curves do not follow complete linear behavior, in the sense that their slope is not constant, but have a curvature depending on K. The smallest curvature has been estimated for the category of toxics and the largest value for the vapor hydrocarbons. This can be explained by the fact that assuming a constant population density in the area of accident and by ignoring all the other factors that may contribute to its severity, toxic substances tend to affect uniformly larger areas ($K = 1.46$), whereas vapor hydrocarbons accidents are expected to be more severe in the vicinity of the accident source ($K = 8.42$). The same procedure is applied for domino accidents. Again the estimated parameters are given in Table 4. Results show clearly that K parameter increases in the case of domino accidents, for all categories, indicating an increased severity of the accident at the direct surrounding area, in the sense that was described above. On the other hand, the α parameter does not seem to follow a certain pattern since it remains the same for the whole set (0.79), and the liquid fuel accidents (0.59) increases for the vapor hydrocarbons (1.06) and decreases for the category of the miscellaneous substances (0.76). No calculations are performed for domino accidents

Table 4
Estimated parameters for Pareto PDF

	α_{all}	K_{all}	α_{domino}	K_{domino}
Sample	0.79	3.54	0.79	4.86
Liquid fuels	0.60	2.18	0.59	2.59
Vapor hydrocarbons	0.88	8.42	1.06	9.90
Toxic substances	0.54	1.46	[a]	[a]
Miscellaneous substances	0.99	3.12	0.73	3.47

[a] Too few data to perform calculations.

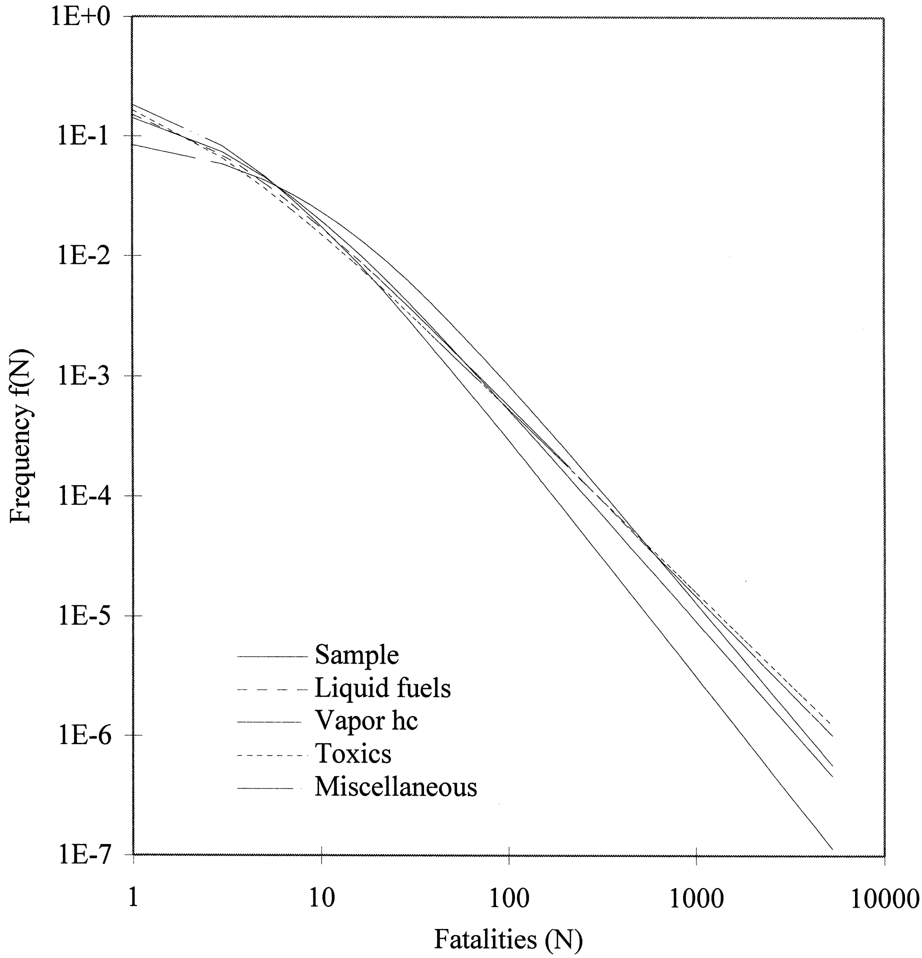

Fig. 3. Theoretical Pareto frequency distributions estimated for each accident category.

involving primarily toxic substances because the number of available accident cases does not allow a significant statistical treatment.

4. Statistical analysis

Chemical accident severity analysis is an active field of research and to date many researchers have focused attention in determining *f–N* curves using historical data and thus calculating the probability of high-consequence accidents happening in the future. Considering that chemical accident behavior does change through time as well as the different data sets used by researchers, Bayesian inference may be used as a method to compare different data sets and modify the distribution parameters as the accident parameters change.

Suppose that $z' = (z_1, \dots, z_n)$ is a vector of n observations whose probability distribution $f(z|\theta)$ depends on the value of k parameters $\theta' = (\theta_1, \dots, \theta_k)$. Suppose also that θ itself has a probability distribution $f(\theta)$. Then, according to Bayes theorem [14]:

$$f(\theta|z) = L(\theta|z) f(\theta) \tag{9}$$

$f(\theta)$ indicates what is known about θ without knowledge of the data and is called the prior distribution of θ, $f(\theta|z)$ indicates what is known about θ given knowledge of the data and is called posterior distribution of θ given z, and $L(\theta|z)$ is the likelihood function of θ for given z. In the case of the Pareto PDF the analysis is performed for the a parameter while K is considered to be constant. The only prior knowledge about α is its estimated mean. It has been proposed by Englehardt [15] that an exponential distribution can be used as a prior for parameter α. The expression for the prior is, then:

$$f(\alpha) = \beta \exp(-\beta\alpha) \tag{10}$$

where β is equal to the inverse of the estimated mean of α whereas the calculated values of α are given in Table 4.

The posterior distribution of α is given by:

$$f(\alpha|z) = \prod_{i=1}^{n} \alpha K^{\alpha} (K + z_i)^{-(\alpha+1)} \beta \exp(-\beta\alpha) \tag{11}$$

or

$$f(\alpha|z_i) = \beta^{n+1} \alpha^n \exp\left[-\alpha\left(\beta + n\overline{\ln(K + z_i)} - n\ln K\right)\right] \tag{12}$$

where $\overline{\ln(K + z_i)}$ is the average of the quantity $\ln(K + z_i)$ and the values of K are also given in Table 4.

The predictive distribution for accident severity, in terms of fatalities, that accounts for uncertainty in the value of α is [16]:

$$f\left(z|K,\beta,n,\overline{\ln(z_i + K)}\right) = \int_0^{\infty} f(z|\alpha) f\left(\alpha|K,\beta,z_1,\dots,z_n\right) \mathrm{d}\alpha. \tag{13}$$

Substituting Eqs. (2) and (12) into the above relation results in the following expression:

$$f\left(z|K,\beta,n,\overline{\ln(z_i + K)}\right) = \frac{(n+1)\left[\beta + n\overline{\ln(K + z_i)} - n\ln K\right]^{n+1}}{(z+k)\left[\ln(z+K) + \beta + n\overline{\ln(z_i + K)} - (n+1)\ln K\right]^{n+2}} \tag{14}$$

where k is a normalization constant.

The corresponding cumulative distribution is:

$$f\left(z|K,\beta,n,\overline{\ln(z_i + K)}\right) = 1 - \left[\frac{\beta + n\overline{\ln(z_i + K)} - n\ln K}{\ln(z+k) + \beta + n\overline{\ln(z_i + K)} - (n+1)\ln K}\right]^{n+1}. \tag{15}$$

The above equations incorporate knowledge of the α parameter variation assuming exponential prior distribution of α.

It is considered that the set of accident cases used for the statistical analysis is a random sample extracted from the accidents population. Then, Eqs. (14) and (15) are expected to be more accurate than Eqs. (2) and (3), respectively. The former incorporate the information taken from the statistically independent, random sample of accidents used, with regard to the probability density of α.

The estimated Pareto and Bayesian cumulative distributions and the available data are plotted, for the range of 1–1000 fatalities, in Fig. 4. The cumulative Bayesian distribution is plotted for $n = 10$ and $n = 100$ observations. It is noticed that as n (the sample population) increases, the Bayesian distribution converges to the estimated Pareto distribution.

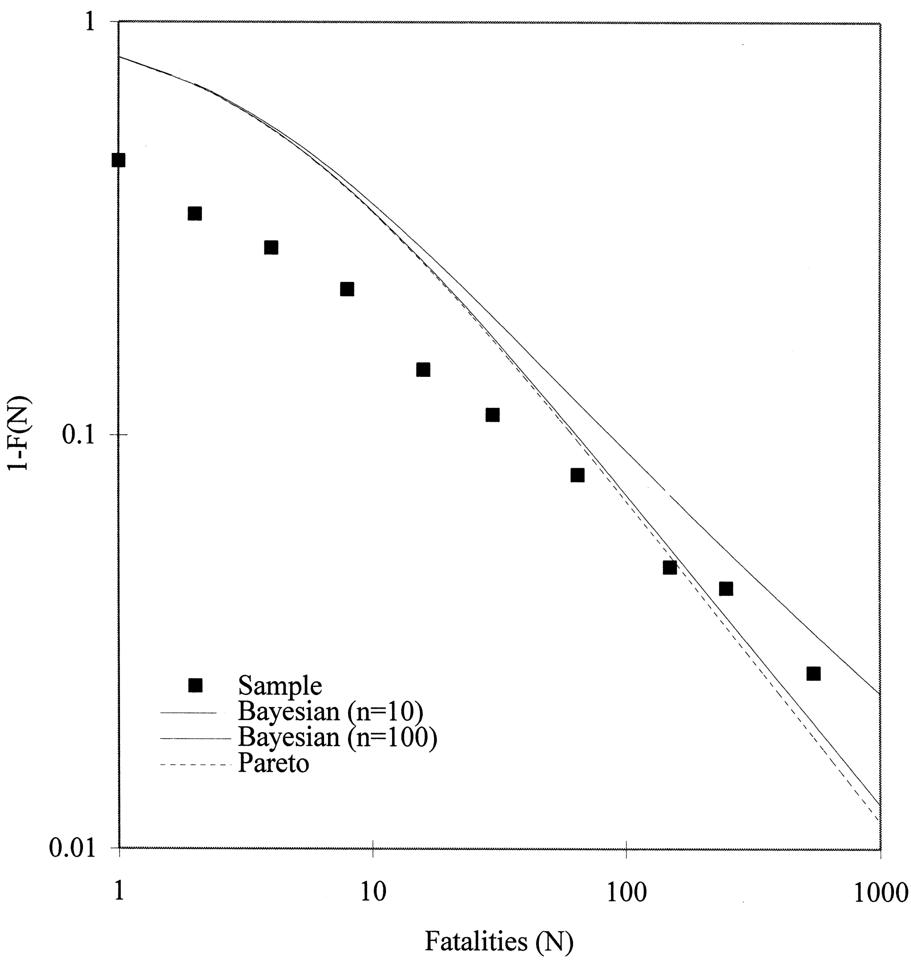

Fig. 4. Comparison of the estimated Pareto and Bayesian $1 - F(N)$ vs. the available data points.

5. Conclusions

In this study, a set of 207 major chemical accidents found in the open literature are statistically analyzed and conclusions are derived regarding the likelihood and the expected severity of domino chemical accidents. It is found that the likelihood of an accident to provoke domino effects depends on the substance involved in it. More specifically, for the set of accidents used, the probability of one domino is approximately 0.39, while accidents involving vapor hydrocarbons are the most likely to cause domino effects, 0.58, followed by accidents involving liquid fuels with a probability of approximately 0.49. On the other hand, accidents involving toxic substances do not usually result in domino effects, 0.16, and the probability of the rest of the accidents (category defined as accidents involving miscellaneous substances), approximately 0.33, is near to that of the whole set.

Differences are also observed, in terms of fatalities caused, depending on the substance involved and whether domino effects are present or not. Domino accidents of the sample are severe and the same holds true for those dominos which involve liquid fuels. However, conclusive results cannot be drawn for accidents involving vapor hydrocarbons and toxic substances by simply inspecting the corresponding f–N curves. The expected severity of a chemical accident is modeled by fitting the fatalities data to a two-parameter modified version of the Pareto distribution. It is found that one of the parameters changes considerably for domino accidents. Comparison of accidents involving vapor hydrocarbons with those involving toxic substances leads to the conclusion that consequences intensity shows a different pattern for each category. Thus, given a constant population density in the area of the accident, consequences due to toxic substances are prone to be dispersed in a wider geographical area.

The theoretical distribution of the accidents severity is further refined by taking into account that its parameters are themselves random variables depending on the random sample of data that has been extracted from the population of historical accidents. This is done with the help of Bayesian analysis. This way, a more representative distribution is derived, which incorporates information with regard to one of the two initial Pareto distribution parameters density, while the second parameter is assumed to be constant. Further analysis should be made taking into account a number of other parameters which are expected to affect the accident behaviour. These parameters may include the quantities of the chemicals involved and population density of the accident area. Bayesian analysis can be used as an inference method in this case as well although complexity will be substantially increased.

Notation

$f()$	Probability density function of ( )
K	Parameter of Pareto distribution
k	Normalization constant

$L()$	Likelihood density of $()$
n	Number of observations of accident size
Z_0	Minimum accident size
z	Accident size
z_i	ith observation of accident size
α	Parameter of Pareto distribution
β	Parameter of exponential prior distribution for Pareto parameter α
θ	Parameter of probability density function

Appendix A. Important chemical accidents of the sample

Country/Year	Substance	Fatalities	Reference
Brazil, Cubato/1984	fuels	508	4
Mexico, Mexico City/1984	LPG	500	4
Soviet Union/1989	LPG	462	4
Spain, San Carlos de la Rapita/1978	propylene	215	4
Germany, Ludwigshafen/1948	dimethyl ether	207	4
UK, Piper Alpha/1988	fuels	167	4
Venezuela, Caracas/1982	fuels	160	4
South Korea, Taegu/1995	LNG	106	9
Ireland, Whiddy Island/1979	light crude oil	50	4
Brazil, Rio de Janeiro/1972	LPG	37	4
USA, Louisiana/1977	cereal dust	36	4
USA, Puerto Rico/1996	propane	33	6
UK, Flixborough/1974	cyclohexane	28	4
USA, Texas/1989	ethylene	23	4
France, Feyzin/1966	LPG	18	4
South Africa, Potchefstrom/1973	ammonia	18	4
Greece, Thriassion/1992	light naphtha	13	7
USA, Tennessee/1978	LPG	16	4
Netherlands, Beek/1975	propylene	14	4
USA, Kentucky/1964	acetylene	12	4
Canada, LaSalle/1966	styrene	11	4
Japan, Sodegaura/1992	hydrogen	10	4
UK, Lincolnshire/1975	metal	11	4
USA, Nebraska/1977	ammonia	8	11
USA, Florida/1978	chlorine	8	11
USA, Texas/1974	vinyl chloride	7	4
Netherlands, Rotterdam/1991	benzoic acid	7	7
USA, Pennsylvania/1995	hydrocarbons	7	5
Belgium, Antwerp/1975	ethylene	6	4
USA, Texas/1976	ammonia	6	4
Italy, Florence/1982	propane	5	7

References

[1] Council Directive 96/82/EC of 9 December 1996 on the control of major-accident hazards involving dangerous substances, Official Journal of the European Communities, Luxembourg, 1997.
[2] G. Apostolakis, Science 250 (1990) 1359.
[3] P. Haastrup, L. Brockhoff, J. Hazard. Mater. 27 (1991) 339.
[4] F.P. Lees, Loss Prevention in the Process Industries, Vol. 3, Butterworth, Heinemann, 1996.
[5] EPA Chemical Accident Investigation Report, Pennzoil Product Company Refinery, Rouseville, PA, Report 550-R-98-001, March 1998.
[6] National Transportation Safety Board, Pipeline Accident Report, PB97-916501, NTSB/PAR-97/01, 1997.
[7] C. Kirchsteiger, Using Modern Database Concepts to Facilitate Exchange of Information on Major Accidents in the European Union, Advances in Safety and Reliability: ESREL'97, 1997, p. 667.
[8] K. Rasmussen, The experience with the major accident reporting system from 1984 to 1993, JRC, Report EUR 16341 EN, 1996.
[9] H. Montiel, J.A. Vilchez, J. Arnaldos, J. Casal, J. Hazard. Mater. 51 (1996) 77.
[10] Health and Safety Executive, Canvey: A Second Report, HMSO, London, 1981.
[11] P. Haastrup, L. Brockhoff, J. Loss Prev. Ind. 3 (1990) 395.
[12] F.P. Lees, Loss Prevention in the Process Industries, Vol. 1, Butterworth, Heinemann, 1996.
[13] R.A. Cooper, A.J. Weekes, Data, Models and Statistical Analysis, Philip Allan, Southampton, 1983.
[14] G.E. Box, G.C. Tiao, Bayesian Inference in Statistical Analysis, Wiley, New York, 1992.
[15] J.D. Englehardt, J. Environ. Eng. 121 (1995) 455.
[16] D.R. Cox, D.V. Hinkley, Theoretical Statistics, Chapman & Hall, London, 1974.