Accident Analysis and Prevention 84 (2015) 20–26

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Analysis of railroad tank car releases using a generalized binomial model

Xiang Liu[a,*], Yili Hong[b]

[a] Department of Civil and Environmental Engineering, Rutgers, The State University of New Jersey, CoRE 606, 96 Frelinghuysen Road, United States
[b] Department of Statistics, Virginia Tech 213 Hutcheson Hall, Blacksburg, VA 24060, United States

ARTICLE INFO

Article history:
Received 3 June 2015
Received in revised form 25 July 2015
Accepted 3 August 2015

Keywords:
Petroleum crude oil
Tank car release
Train accident
Generalized binomial model
Probability

ABSTRACT

The United States is experiencing an unprecedented boom in shale oil production, leading to a dramatic growth in petroleum crude oil traffic by rail. In 2014, U.S. railroads carried over 500,000 tank carloads of petroleum crude oil, up from 9500 in 2008 (a 5300% increase). In light of continual growth in crude oil by rail, there is an urgent national need to manage this emerging risk. This need has been underscored in the wake of several recent crude oil release incidents. In contrast to highway transport, which usually involves a tank trailer, a crude oil train can carry a large number of tank cars, having the potential for a large, multiple-tank-car release incident. Previous studies exclusively assumed that railroad tank car releases in the same train accident are mutually independent, thereby estimating the number of tank cars releasing given the total number of tank cars derailed based on a binomial model. This paper specifically accounts for dependent tank car releases within a train accident. We estimate the number of tank cars releasing given the number of tank cars derailed based on a generalized binomial model. The generalized binomial model provides a significantly better description for the empirical tank car accident data through our numerical case study. This research aims to provide a new methodology and new insights regarding the further development of risk management strategies for improving railroad crude oil transportation safety.

1. Introduction

The United States is experiencing an unprecedented boom in the production of petroleum crude oil and natural gases from shale, driven by technological advances in hydraulic fracturing and horizontal drilling. This has consequently led to a significant rise in the rail transport of petroleum crude oil. In 2014, there were over 500,000 carloads of petroleum crude oil transported over U.S. rail network, up from 9500 in 2008, or a 5300% increase (Barkan et al., 2015) (Fig. 1). The fast-growing crude oil traffic puts railroad safety under the national spotlight, especially in the wake of a chain of crude oil release incidents in 2013 and 2014, such as those in Lac-Mégantic, Canada in July 2013; Aliceville, Alabama in November 2013; Casselton, North Dakota in December 2013; and Lynchburg, Virginia in April 2014.

Differing from roadway transport of hazardous materials, which usually involves a single tank trailer, a train may carry multiple tank cars loaded with hazardous materials. In particular, railroads promote the use of unit-trains (a unit-train may contain 50 to over 100 loaded hazardous materials cars) to transport crude oil as a means of achieving greater transportation efficiency (AAR, 2015). In view of continual growth in railroad crude oil traffic, a proper risk management becomes a high priority for railroad carriers, regulators, shippers and related stakeholders (CRS, 2014).

In railroad hazardous materials transportation risk analyses, the number of tank cars releasing per train accident is an important variable (Glickman et al., 2007; Bagheri et al., 2011, 2012, 2014; Liu et al., 2014). Previous studies were almost exclusively based upon the same assumption that tank car releases in the same train accident are mutually independent. Based on this assumption, a binomial (or Poisson binomial) model was developed to estimate the number of tank cars releasing per accident, given the total number of tank cars derailed. However, it is possible that there exists interdependency between tank car releases within the same train accident, due to the interactive effects of derailed tank cars that are coupled together in the same block, or due to certain common accident conditions. If such dependency exists, it will affect the estimation of the probability of a large, multiple-car release incident. To our knowledge, there is no prior research that specifically focused on modeling dependent tank car releases within the same accident.

* Corresponding author.
E-mail addresses: xiang.liu@rutgers.edu (X. Liu), yilihong@vt.edu (Y. Hong).

http://dx.doi.org/10.1016/j.aap.2015.08.004
0001-4575/

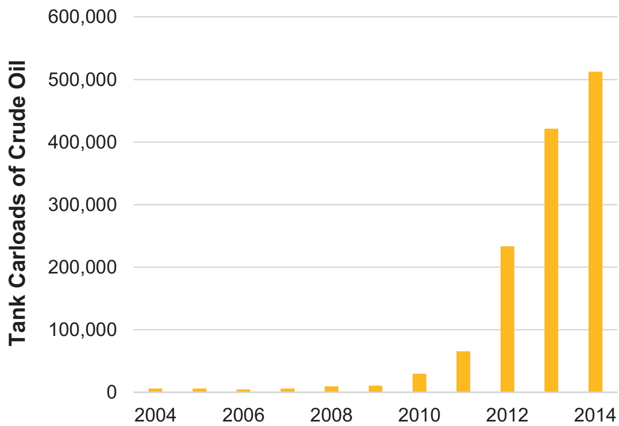

Fig. 1. Traffic of rail transport of petroleum crude oil on U.S. Class I railroads. *Source*: Association of American Railroads, adapted from Barkan et al. (2015).

To narrow this knowledge gap, we first formulate each railroad tank car release as a Bernoulli variable and analyze their sum using a generalized binomial model. Next, we present a numerical example to illustrate the application of the new model, with comparison to the previous binomial model. Finally, we discuss the implications of this study with respect to rail safety policy and practice. The methodology developed in this paper can be expanded to a larger risk management framework for improving the safety of rail transport of crude oil and other hazardous materials.

2. Literature review

Each derailed tank car containing hazardous materials has one of two outcomes: release or no release. The release of each derailed tank car can be viewed as a Bernoulli variable, whose Bernoulli probability is referred to as the conditional probability of release (CPR) (Barkan et al., 2007; Barkan, 2008). The total number of tank car releases in a train accident can be viewed as the sum of a series of Bernoulli variables. If the Bernoulli variables are independent and identically distributed (IID), their sum follows a binomial distribution (Ross, 2007). Previous studies almost exclusively used a binomial model to estimate the number of tank cars releasing given the number of tank cars derailed (e.g., Nayak et al., 1983; Glickman et al., 2007; Bagheri et al., 2011, 2012, 2014). Liu et al. (2014) extended the prior work by accounting for heterogeneous, independent tank car releases based on a Poisson binomial model.

However, we are unaware of any published study yet to consider the total number of *dependent* tank car releases in a train accident. This paper aims to narrow this knowledge gap by proposing a new generalized binomial model.

3. Methodology

The release of each derailed tank car is a Bernoulli variable. Let D_i denote the release of the ith derailed tank car in a train. The Bernoulli probability (denoted as P_i) measures tank car safety performance in an accident (Barkan et al., 2007; Barkan, 2008) (Table 1).

Table 1
Outcomes of a derailed tank car in a train accident.

	Bernoulli variable (D_i)	Probability
Release	1	P_i
No release	0	$1-P_i$

If there are n tank cars derailed in a train accident, the total number of tank cars releasing (denoted as R) is expressed as:

$$R=\sum_{i=1}^{n} D_i \tag{1}$$

If the releases of individual derailed tank cars are independent and identically distributed, the total number of releases, R, follows a binomial distribution (Nayak et al., 1983; Glickman et al., 2007; Bagheri et al., 2011, 2012, 2014). When the D_i's are non-identical but independent, R follows a Poisson binomial distribution (Liu et al., 2014). In this paper, we analyze the probability distribution of total number of dependent tank car releases, given the total number of tank cars derailed.

First, we define the dependencies among tank car releases within the same accident. We assume that the release probability of the ith derailed tank car within all tank cars derailed depends on the total number of cars releasing prior to it (Eq. (2)):

$$P(D_i=1|D_1,D_2,\ldots,D_{i-1})=P(D_i=1|D_1+D_2+\cdots+D_{i-1}) \tag{2}$$

This assumption indicates that whether a derailed tank car releases depends on the total number of tank cars releasing prior to the car in question, regardless of which of the prior cars have released. This total-kinetic-energy-related assumption appears to agree with the prior railroad engineering research (Barkan et al., 2003; Liu et al., 2011, 2012). For all the tank cars derailing prior to the ith car ($D_1, D_2, \ldots, D_{i-1}$), the total number of their releases would range from 0 to $i-1$. For illustration simplicity, we use the following notation, originally developed by Yu and Zelterman (2002):

$$C_n(s)=P(D_1=1|D_1+D_2+\cdots+D_{n-1}=s-1 \tag{3}$$

where $C_n(s)$ = the conditional probability that the nth derailed tank car would release, given that there are $(s-1)$ tank cars releasing prior to it ($s\geq 1$).

We also define that $C_1(1)=P(D_1=1)$. Let $P_n(s)$ represent the probability of releasing s tank cars out of n derailed tank cars, that is:

$$P_n(s)=P(D_1+\cdots+D_n=s) \tag{4}$$

Based on the Law of Total Probability (Ross, 2007), we have

$$\begin{aligned}P_n(S)=&\ C_n(s-1)P(D_1+\cdots+D_{n-1}=s-1)\\&+[1-C_n(s)]P(D_1+\cdots+D_{n-1}=s)\end{aligned} \tag{5}$$

Eq. (5) can be re-written in an equivalent but more concise way:

$$P_n(S)=C_n(s-1)P_{n-1}(s-1)+[1-C_n(s)]P_{n-1}(s) \tag{6}$$

Eq. (6) provides a recursive algorithm to calculate the probability mass function (PMF) of the number of dependent tank car releases (s) given the number of tank cars derailed (n). If $n-1<s$, $P_{n-1}(s)$ is 0. $P_n(0)$ can be calculated as a complementary probability of $P_n(s\geq 1)$. To explicitly describe the dependency structure among Bernoulli variables, Yu and Zelterman (2002) proposed the following $C_n(s)$ based on a medical research study. In the context of railroad safety, this particular type of dependency structure takes into account the number of tank cars derailed, which is a proxy related to train accident kinetic force and, correspondingly, to the degree of severity (Barkan et al., 2003). Therefore, this paper starts with this dependency structure. In future research, the recursive algorithm proposed in Eq. (6) can be adapted to other possible dependency structures:

$$C_n(s)=\frac{\alpha s+p}{n\alpha+1} \tag{7}$$

Table 2
Illustrative comparison of generalized binomial versus binomial distributions.

Number of tank cars releasing (s)	Probability in the generalized binomial model	Probability in the binomial model	Percent difference
(a) Generalized binomial ($n=4$; $p=0.6$; $\alpha=0.05$); binomial ($n=4$; $p=0.6$)			
0	0.043	0.026	69.8%
1	0.179	0.154	16.8%
2	0.328	0.346	−5.2%
3	0.315	0.346	−8.9%
4	0.135	0.130	4.1%
(b) Generalized binomial ($n=4$; $p=0.6$; $\alpha=-0.05$); binomial ($n=4$; $p=0.6$)			
0	0.010	0.026	−61.7%
1	0.114	0.154	−25.8%
2	0.364	0.346	5.3%
3	0.391	0.346	13.1%
4	0.121	0.130	−6.4%

where $0<p<1$ and $\alpha \geq -p/N$ (where N is the largest number of tank cars derailed per accident). Based on Eqs. (6) and (7), we can derive the probability distribution of dependent tank cars releasing in the same accident. The mean and variance of the number of tank cars releasing are:

$$E[R] = np \tag{8}$$

$$Var[R] = np(1-p)\{1+(n-1)\alpha/(\alpha+1)\} \tag{9}$$

The covariance between any two tank car releases in a train accident is

$$Cov[D_i, D_j] = \alpha p(1-p)/(a+1) \tag{10}$$

Eq. (8) shows that the generalized binomial model (based on a particular dependency structure) has the same mean value with the binomial model. However, the generalized binomial model has a different variance function (Eq. (9)). In Eq. (10), α indicates the sign of the correlation between tank car releases. If $\alpha=0$, the generalized binomial model is identical to the binomial model. If $\alpha>0$, tank car releases within the same accident are positively correlated, and the generalized binomial model will allow for a greater variance than the binomial model. Otherwise, if $\alpha<0$, tank car releases are negatively correlated, and the generalized binomial model would have a smaller variance than the binomial model. Eq. (10) shows that the value of parameter α also indicates the degree of dependency. The larger the value of α, the larger the covariance between tank car releases. To better understand the effect of the dependency factor α, consider the following hypothetical example in which $n=4$ (4 tank cars derailed in an accident). In the binomial distribution, it is assumed that p = 0.60. For comparison, in the generalized binomial distribution, it is assume that $p=0.60$, $\alpha=0.05$(or -0.05). Using the recursive algorithm introduced in Eq. (6), we calculated the probability distribution of the number of tank cars releasing in the binomial distribution versus in the generalized binomial distribution (Table 2) (the calculation details are in Appendix A).

Table 3 shows that if tank car releases within the same train accident are positively correlated (in the hypothetical example, $\alpha=0.05$), the generalized binomial model will have a higher probability estimation on a no-release or small release incident. The generalized binomial model also estimates a higher probability for an all-tank-car-release incident. It indicates that the derailed tank cars tend to release (or do not release) together, when their release probabilities are positively correlated. By contrast, if tank car releases within the same train accident are negatively correlated (in the hypothetical example, $\alpha=-0.05$), the generalized model will result in a lower estimation on a small release (or no release) as well as on an all-car release incident.

Table 3
Parameter estimates based on a sample of train accident data.

	Binomial	Generalized binomial
Average release probability (p)	0.2162	0.2223
Correlation parameter (α)	0	0.6558
Likelihood	7.80E−22	1.13E−19
Log-likelihood	−48.60	−43.62

In the generalized binomial model, the parameter p is interpreted as the average release probability of a derailed tank car; the parameter α represents the correlation parameter between tank car releases. Both parameters can be fitted to the empirical data using the method of maximum likelihood (ML):

$$(p,\alpha)_{ML} = arg_{(p,\alpha)} \max \prod_k \ln[P(R_k)] \tag{11}$$

where (p,α) = parameters in the generalized binomial model and $P(R_k)$ = probability of releasing R_k tank cars in the kth train accident.

4. Numerical example

4.1. Dataset

To further illustrate our methodology, we assembled an anonymous sample of 56 hazardous-materials train accidents from 1990 to 2010. The data used for the statistical analysis are in Appendix B. In those train accidents, all the derailed tank cars conform to non-jacketed DOT 111A100W1 tank car design features (e.g., 7/16 in. tank thickness). This was one of common hazardous materials tank cars used in North America. The derailment speed is around 30 mph. Each train accident resulted in 10 railcars derailed (including both hazmat cars and non-hazmat cars). The selection of "homogeneous" accident conditions could better isolate the effect of tank car release dependency by controlling other factors. Based on Eq. (11), we fitted the empirical data using:

(1) Binomial model (previous model). The average release probability was estimated based on the empirical data. The correlation parameter was assumed to be zero.
(2) Generalized binomial model (new model). Both the average release probability and correlation parameter were estimated based on the empirical data.

4.2. Parameter estimates

In this paper, we used the ML method to determine the estimates of the unknown parameters. Because there is no closed-form expression for the ML estimates (p,α), a numeric method was applied based on the optim() function in software R. The results of parameter estimation are in Table 3.

The binomial distribution model has a single parameter p (average release probability of a derailed tank car), whereas the generalized binomial model has two parameters, which are the average release probability (p) and the correlation (α) between the releases of tank cars in the same train accident. There are several observations:

- The parameter p is very close in both models. This is not surprising because both models have the same mathematical expression of the mean (Eq. (8) of this paper).
- The real difference between the two models is that the generalized binomial model accounts for the dependencies among tank car releases in the same train accident. By contrast, the binomial model assumes independency.

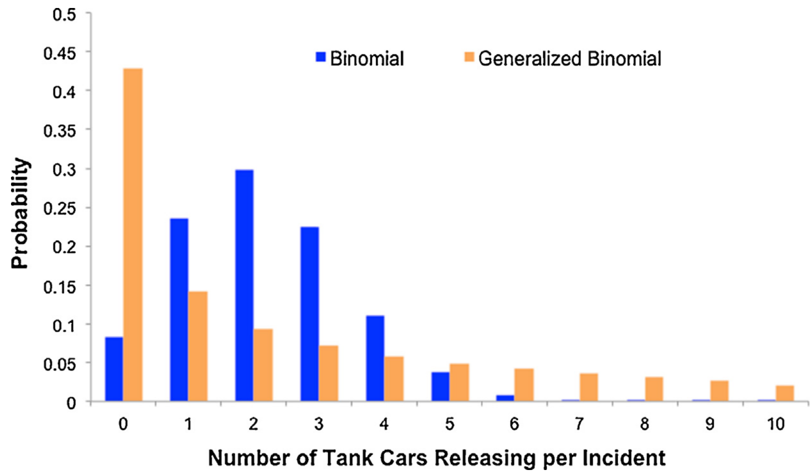

Fig. 2. Distributions of tank cars releasing by different models (10 tank cars derailed per accident). Notes: (1) In the binomial model, $p = 0.2162$. (2) In the generalized binomial model, $p = 0.2223$, $\alpha = 0.6558$ based on the sample data described in Section 4.

Each model's goodness-of-fit is evaluated based on the likelihood value at the ML estimate (or the corresponding log-likelihood value). A higher likelihood value indicates a better fit with the empirical data. As shown in Table 3, the likelihood of the generalized binomial model (1.13E−19) is significantly higher than the binomial model (7.80E−22). This distinction indicates that our sample train accident dataset exhibits an interdependency among tank car releases that is not captured by the binomial model. To further validate this conclusion, we conducted a likelihood ratio (LR) test (Agresti, 2007) to examine whether the correlation parameter, α, is zero.

Null Hypothesis H_0: $\alpha = 0$ (no tank car release dependency within the same train accident).
Alternative Hypothesis H_a: α is not equal to 0.

To test the hypothesis, a statistic called Deviance is calculated as follows:

$$D = 2 \times [\ln L(\text{generalized binomial}) - \ln L(\text{binomial})] \quad (12)$$

where D = deviance, $\ln L$(generalized binomial) = logarithmic likelihood of a generalized binomial model and $\ln L$(binomial) = logarithmic likelihood of a binomial model.

According to the statistical theory, the deviance approximately follows a chi-square distribution (Agresti, 2007). Using the information from Table 3, $D = 2 \times [-43.62 - (-48.60)] = 9.96$. The corresponding P-value is 0.002 (degree of freedom is one), indicating that the null hypothesis is rejected. Therefore, the correlation parameter (α) is significantly different from zero. Therefore, the inclusion of an additional correlation parameter better fits the empirical dataset. Due to information limit, we are unable to duplicate the analysis for all historical tank car accidents. For the particular sample data used in this paper, the generalized binomial model shows a better fit and appears to be a promising approach to estimating the total number of dependent tank car releases.

The LR test above shows that the generalized binomial model outperforms the binomial model, based on the sample data. Next, we conduct a Chi-square goodness-of-fit test to study whether the generalized binomial model adequately fits the empirical data, using the following equation:

$$x^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \quad (13)$$

where O_i = observed number of tank cars releasing per accident and E_i = expected number of tank cars releasing per accident.

Using the generalized binomial (Table 3) and raw data (Appendix B), $\chi^2 = 61.2$. The corresponding P-value is 0.31 (degree of freedom = 55). These outcomes show that the generalized binomial model can adequately fit the empirical data used in this paper.

5. Discussion

In this section, we discuss the implications of our research with respect to rail safety policy and practice. The generalized binomial model appears to be a promising alternative to the previous binomial model for analyzing the number of railroad tank car releases. Statistically, a generalized binomial model can be viewed as the sum of *dependent* Bernoulli variables, whose special case is the binomial model when the dependency is zero. If the dependency exists, it is important to understand how the use of different models (generalized binomial versus binomial) can affect risk analysis. To illustrate, consider the following example.

Given that a train derailment results in 10 tank cars derailed, what is the probability distribution of the number of tank cars releasing (ranging from 0 to 10)? Using the parameter estimates developed in the previous section (Table 3), the generalized binomial model predicts a higher probability for either no-release or a large release (five or more cars releasing). By contrast, the binomial model predicts a higher probability on other release magnitudes (Fig. 2).

Similarly, we consider the distributions of the number of tank cars releasing, given 5, 15, 20, and 25 tank cars derailed, respectively, using the same fitted parameters (Fig. 3).

The spate of recent crude oil release incidents has raised interest in better understanding the likelihood of a large, multiple-car release incident. This paper shows that, if there are positive correlations between tank car releases within the same train accident, the binomial model may underestimate the probability of a large, multiple-car release incident, when compared with the generalized binomial model. Therefore, when the interdependency exists, we propose consideration of the generalized binomial model to better understand the risk of a large, multiple-car release incident. The magnitude of the dependency may depend on both train operating factors and accident characteristics, which might be specific to different datasets. In this paper, we conduct a proof-of-concept pilot study based on a specific sample dataset. The data limit constrains us to perform a nationwide analysis at this moment. Also, the analysis of various other tank car release dependency structures using

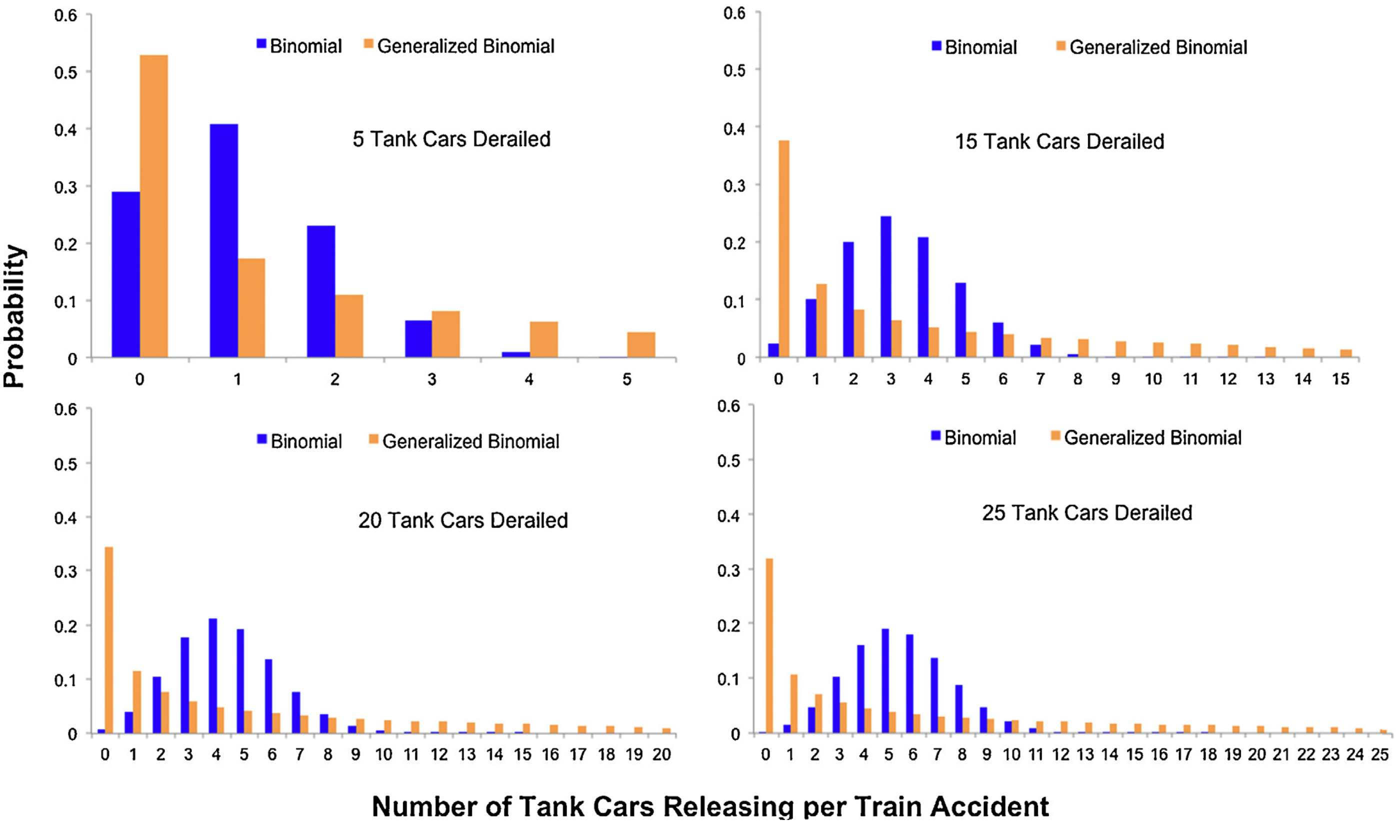

Fig. 3. Model comparison for other derailment severity scenarios. Notes: (1) In the binomial model, $p = 0.2162$. (2) In the generalized binomial model, $p = 0.2223$, $\alpha = 0.6558$ based on the sample data described in Section 4.

other datasets is left for more extensive longer-term research, as that information becomes available.

This paper focuses on releases from derailed tank cars, without accounting for possible releases from non-derailed cars due to thermal tear in a fire event. The fire impinging on the tank would weaken the steel on the upper side, reduce its strength and possibly cause a tank car release (Barkan et al., 2015). Industry and government had sponsored the development of an engineering tool known as the Analysis of Fire Effects on Tank Cars (AFFTAC) to evaluate increases in thermally induced pressure and the effectiveness of designs for pressure relief devices (Runnels, 2013; Barkan et al., 2015). Future research could be developed to incorporate fire-caused tank car releases into railroad risk analysis modeling. Also, a large body of the literature assumes that train accidents are independent. Future research might be needed to explore the scenarios in which inter-dependencies among train accidents exist as well, in addition to accounting for inter-dependency between tank car releases within the same train accident.

6. Conclusion

This paper focuses on modeling the probability distribution of the number of dependent tank car releases within the same train accident, through a proof-of-concept study based on a sample train safety data. A generalized binomial model is developed in order to analyze the total number of dependent railroad tank car releases, given the number of tank cars derailed. The generalized binomial model exhibits a better fit to the empirical sample data than the binomial model. In view of growing national interest in managing the risk of the rail transport of petroleum crude oil and other types of chemicals, this research can provide new methods to improve the accuracy and reliability of railroad transportation risk analysis. In the next step, the methodology developed in this paper will be expanded to a larger integrated risk management framework for improving the safety of rail transport of hazardous materials.

Acknowledgements

The authors thank Dr. Chris Barkan, Mr. Todd Treichel and Dr. Rapik Saat for their comments on the earlier draft. However, the authors are solely responsible for all the views and analyses herein.

Appendix A. Example calculation of recursive algorithm

It is assumed that a binomial distribution ($n = 4$; $p = 0.6$) and a generalized binomial distribution ($n = 4$; $p = 0.6$; $\alpha = 0.05$). The probability mass function (PMF) of a generalized binomial distribution is estimated using the recursive algorithm described in Eq. (6). The section below shows a step-by-step manual calculation process for illustration purpose. A computer program is developed to automate all the analyses presented in this paper.

Step 1: By definition, $P_1(0) = 0.4$; $P_1(1) = 0.6$.
Step 2: Using Eq. (7), $C_2(0) = 0.545$; $C_2(1) = 0.590$; $C_2(2) = 0.636$. Using Eq. (6), $P_2(1) = C_2(0)P_1(0) + [1 - C_2(1)]P_1(1) = 0.464$. Similarly, $P_2(2) = C_2(1)P_1(1) + [1 - C_2(2)]P_1(2) = 0.355$. Note that, by definition, $P_1(2) = 0$. Finally, $P_2(0) = 1 - P_2(1) - P_2(2) = 0.182$.
Step 3: Repeat the procedure to estimate $P_3(0)$, $P_3(1)$, $P_3(2)$, $P_3(3)$ based on $P_2(0)$, $P_2(1)$, $P_2(2)$.
Step 4: Repeat the procedure to estimate $P_4(0)$, $P_4(1)$, $P_4(2)$, $P_4(3)$, $P_4(4)$ based on $P_3(0)$, $P_3(1)$, $P_3(2)$, $P_3(3)$.

Appendix B. Tank car release dataset and results

ID	Number of tank cars derailed	Number of tank cars releasing	Estimated probability using binomial model	Estimated probability using generalized binomial model	Difference in estimated probability
1	1	0	0.78	0.78	−0.29%
2	1	1	0.22	0.22	1.05%
3	1	0	0.78	0.78	−0.29%
4	1	1	0.22	0.22	1.05%
5	2	0	0.61	0.67	10.67%
6	4	0	0.37	0.56	52.09%
7	1	0	0.78	0.78	−0.29%
8	5	0	0.29	0.53	83.02%
9	1	0	0.78	0.78	−0.29%
10	3	0	0.47	0.61	28.23%
11	2	1	0.34	0.21	−39.15%
12	2	0	0.61	0.67	10.67%
13	1	0	0.78	0.78	−0.29%
14	1	0	0.78	0.78	−0.29%
15	1	0	0.78	0.78	−0.29%
16	6	4	0.02	0.06	191.25%
17	2	0	0.61	0.67	10.67%
18	1	0	0.78	0.78	−0.29%
19	2	0	0.61	0.67	10.67%
20	2	0	0.61	0.67	10.67%
21	1	1	0.22	0.22	1.05%
22	1	0	0.78	0.78	−0.29%
23	3	0	0.47	0.61	28.23%
24	3	0	0.47	0.61	28.23%
25	1	0	0.78	0.78	−0.29%
26	1	0	0.78	0.78	−0.29%
27	1	1	0.22	0.22	1.05%
28	1	0	0.78	0.78	−0.29%
29	1	0	0.78	0.78	−0.29%
30	3	0	0.47	0.61	28.23%
31	1	0	0.78	0.78	−0.29%
32	1	1	0.22	0.22	1.05%
33	1	0	0.78	0.78	−0.29%
34	1	0	0.78	0.78	−0.29%
35	2	1	0.34	0.21	−39.15%
36	2	1	0.34	0.21	−39.15%
37	2	1	0.34	0.21	−39.15%
38	4	1	0.42	0.18	−56.33%
39	2	0	0.61	0.67	10.67%
40	3	3	0.01	0.08	634.67%
41	2	1	0.34	0.21	−39.15%
42	1	0	0.78	0.78	−0.29%
43	1	0	0.78	0.78	−0.29%
44	1	1	0.22	0.22	1.05%
45	2	2	0.05	0.12	143.57%
46	1	0	0.78	0.78	−0.29%
47	6	0	0.23	0.50	122.45%
48	1	0	0.78	0.78	−0.29%
49	4	2	0.18	0.11	−34.93%
50	3	0	0.47	0.61	28.23%
51	1	0	0.78	0.78	−0.29%
52	3	0	0.47	0.61	28.23%
53	1	0	0.78	0.78	−0.29%
54	1	1	0.22	0.22	1.05%
55	4	0	0.37	0.56	52.09%
56	4	0	0.37	0.56	52.09%
Total	111	24			
Likelihood			7.80E−22	1.13E−19	
Log-likelihood			−48.60	−43.62	

Note: In the binomial distribution, the maximum likelihood (ML) estimator of tank car probability of release is 0.2162 (24/111 = 0.2162). Probability estimates were rounded.

References

Agresti, A., 2007. Introduction to Categorical Data Analysis, 2 ed. Wiley-Interscience.

Association of American Railroads (AAR), 2015. U.S. Rail Crude Oil Traffic, https://www.aar.org/BackgroundPapers/US Rail Crude Oil Traffic (June 2015).pdf.

Barkan, C.P.L., Ukkusuri, S., Waller, S.T., 2007. Optimizing the design of railway tank cars to minimize accident-caused releases. Comput. Oper. Res. 34, 1266–1286.

Barkan, C.P.L., Dick, C.T., Anderson, R.T., 2003. Analysis of railroad derailment factors affecting hazardous materials transportation risk. Transp. Res. Rec. 1825, 64–74.

Barkan, C.P.L., 2008. Improving the design of higher-capacity railway tank cars for hazardous materials transport: optimizing the trade-off between weight and safety. J. Hazard. Mater. 160, 122–134.

Barkan, C.P.L., Liu, X., Saat, M.R., 2015. Enhanced tank car design improves the safety of transporting crude oil and alcohol by rail. TR News 298, 41–43.

Bagheri, M., Saccomanno, F.F., Chenouri, S., Fu, L.P., 2011. Reducing the threat of in-transit derailments involving dangerous goods through effective placement along the train consist. Accid. Anal. Prev. 43, 613–620.

Bagheri, M., Saccomanno, F.F., Chenouri, S., Fu, L.P., 2012. Modelling hazardous materials risks for different train make-up plans. Transp. Res. Part E: Logist. Transp. Rev. 48, 907–918.

Bagheri, M., Verma, M., Verter, V., 2014. Transport mode selection for toxic gases: rail or road? Risk Anal. 34 (1), 168–186.
Glickman, T.S., Erkut, E., Zschocke, M.S., 2007. The cost and risk impacts of rerouting railroad shipments of hazardous materials. Accid. Anal. Prev. 39 (5), 1015–1025.
Congressional Research Service (CRS), 2014. U.S. Rail Transportation of Crude Oil: Background and Issues for Congress.
Liu, X., Barkan, C.P.L., Saat, M.R., 2011. Analysis of derailments by accident cause: evaluating railroad track upgrades to reduce transportation risk. Transp. Res. Rec. 2261, 178–185.
Liu, X., Saat, M.R., Barkan, C.P.L., 2012. Analysis of causes of major train derailment and their effect on accident rates. Transp. Res. Rec. 2289, 154–163.
Liu, X., Saat, M.R., Barkan, C.P.L., 2014. Probability analysis of multiple-tank-car release incidents in railway hazardous materials transportation. J. Hazard. Mater. 276, 442–451.
Nayak, P.R., Rosenfield, D.B., Hagopian, J.H., 1983. Event Probabilities and Impact Zones for Hazardous Materials Accidents on Railroads, Report DOT/FRA/ORD-83/20, FRA. U.S. Department of Transportation, Federal Railroad Administration, Washington, DC.
Ross, S., 2007. Introduction to Probability Models. Academic Press.
Runnels, S., 2013. AFFTAC Analysis of Fire Effects on Tank Cars. RSI-AAR Railroad Tank Car Safety Research and Test Project.
Yu, C., Zelterman, D., 2002. Sums of dependent Bernoulli random variables and disease clustering. Stat. Prob. Lett. 57, 363–373.