Accident Analysis and Prevention 107 (2017) 164–172

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Full length article

Analysis of multiple tank car releases in train accidents

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

[a] Department of Civil and Environmental Engineering, Rutgers University, United States
[b] Department of Statistics, Rutgers University, United States
[c] Department of Statistics, Virginia Tech, United States

ARTICLE INFO

Keywords:
Hazardous materials
Safety
Tank car release
Correlated binomial model
Transportation risk

ABSTRACT

There are annually over two million carloads of hazardous materials transported by rail in the United States. The American railroads use large blocks of tank cars to transport petroleum crude oil and other flammable liquids from production to consumption sites. Being different from roadway transport of hazardous materials, a train accident can potentially result in the derailment and release of multiple tank cars, which may result in significant consequences. The prior literature predominantly assumes that the occurrence of multiple tank car releases in a train accident is a series of independent Bernoulli processes, and thus uses the binomial distribution to estimate the total number of tank car releases given the number of tank cars derailing or damaged. This paper shows that the traditional binomial model can incorrectly estimate multiple tank car release probability by magnitudes in certain circumstances, thereby significantly affecting railroad safety and risk analysis. To bridge this knowledge gap, this paper proposes a novel, alternative Correlated Binomial (CB) model that accounts for the possible correlations of multiple tank car releases in the same train. We test three distinct correlation structures in the CB model, and find that they all outperform the conventional binomial model based on empirical tank car accident data. The analysis shows that considering tank car release correlations would result in a significantly improved fit of the empirical data than otherwise. Consequently, it is prudent to consider alternative modeling techniques when analyzing the probability of multiple tank car releases in railroad accidents.

1. Introduction

Each year, over two million carloads of hazardous materials (hazmat) are transported by American railroads (AAR, 2017). Although hazardous materials accounts for only 7% of U.S. rail traffic, it is responsible for a major share of railroads' liability and insurance risk (AAR, 2017). Since 2005, the shale oil production boom in North America has led to significant growth in rail transport of flammable liquids. Being different from roadway transport of hazardous materials, a train can carry multiple tank cars, sometimes over 100 tank cars in a single train. Therefore, a train accident has the potential to cause the derailments and releases of multiple tank cars. Several recent multiple-tank-car release incidents, particularly the derailments in Lac-Mégantic, Canada in July 2013, Aliceville, Alabama in November 2013, and Casselton, North Dakota in December 2013, all underscore the vital importance of understanding and preventing multiple-car release risk (Liu et al., 2014; Liu, 2017).

One principal task in railroad hazmat transportation risk management is to understand the number of tank cars releasing per train accident. Previous studies predominantly assumed that tank car releases per train accident are mutually independent. Under this assumption, binomial distribution has been used to estimate the number of tank cars releasing given the total 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). To our knowledge, Liu and Hong (2015) was the only published study that accounts for the dependency between tank car releases in the same accident. They found that Beta Binomial model outperforms the traditional binomial model based on one empirical dataset. Their study finds that accounting for tank car release dependency could substantially change risk estimation for the incidents involving a large number of tank cars releasing contents. Therefore, an accurate estimation of multiple tank car release probability is very critical for railroad hazardous materials risk management.

However, Liu and Hong (2015) paper has two major limitations. First, only one type of dependency structure is considered. It is worth investigating whether other dependency structures could further improve the fit of the empirical data. Second, they focused on modeling the conditional mean value of the number of tank cars releasing per accident. In addition to the conditional mean, other distributional statistics (e.g. median, 80th percentile) are also worth investigation,

* Corresponding author.
E-mail address: xiang.liu@rutgers.edu (X. Liu).

http://dx.doi.org/10.1016/j.aap.2017.07.007
Received 16 May 2017; Received in revised form 5 July 2017; Accepted 5 July 2017

especially when the distribution of tank cars releasing is asymmetrical.

This paper aims to advance railroad hazardous materials transportation risk analysis, with the following two objectives:

- Modeling multiple tank car release probability using alternative correlated binomial models (including Beta Binomial model, Increment Risk model and Family History model, respectively), in comparison with the binomial model that assumes no dependency between tank car releases
- Explore the use of quantile statistics to measure the severity of a railroad hazmat release incident, in addition to using the conditional mean

This research focuses on the releases caused by mechanical damage incurred by tank cars in train accidents, without accounting for the releases resulting from thermal tearing, which is a process by which a fire impinging on the tank causes the steel to weaken. Accounting for thermal-tearing-caused tank car release risk is the next step of this work.

The paper is structured as follows. Section 2 presents a review of the literature and clarifies the intended contributions of this paper. Section 3 introduces the statistical methodology that is comprised of three types of correlated binomial models. Section 4 presents the data used for statistical modeling. Sections 5 to 7 discuss the results and implications to railroad safety analysis. Sections 8 and 9 conclude the study and suggest possible future research directions.

2. Literature review

Tank cars today are the second most common type of railroad freight car in North America, accounting for approximately 20 percent of the rail car fleet (Barkan et al., 2013). Tank cars annually transport over two million shipments of hazardous materials that are essential to the nation's economy (Barkan et al., 2013).

The Railway Supply Institute (RSI) and the Association of American Railroads (AAR) developed industry-wide tank car accident statistics since the 1970s (Treichel et al., 2006). Using this database, the AAR-RSI published statistics regarding the safety performance of a tank car by its safety design. For example, if a non-jacketed 111A100W1 (7/16 inch tank thickness) derails, its release probability is 0.196. By contrast, the release probability of a jacketed CPC-1232 car (7/16 inch tank thickness) is reduced to 0.046. Note that the published AAR-RSI tank car accident statistics focus on single tank cars, without accounting for the possible correlation between multiple tank car releases within the same train accident.

In railroad hazmat transportation risk analysis, the estimation of the number of tank cars releasing is a pivotal task. Differing from roadway transport of hazardous materials, a train accident can potentially cause the derailment and releases of multiple tank cars. Given the total number of tank cars derailed, the number of tank cars releasing hazardous materials follows a probabilistic distribution, depending on whether tank car releases within the same accident are independent:

a. Derailed tank cars have independent release probabilities. Almost all previous studies were based on this assumption and they used a binomial distribution to estimate the total number of tank car releases given the number of tank cars derailed (e.g. Nayak et al., 1983; Glickman et al., 2007; Bagheri et al., 2011, 2012, 2014).
b. Derailed tank cars in the same accident have correlated release probabilities. This scenario accounts for the interactions among tank car release probabilities within the same train accident. To our knowledge, the only published study addressing this scenario was presented by Liu and Hong (2015). They used a Beta Binomial model to describe a specific correlation structure between releasing tank cars, and found that the Beta Binomial model outperformed the traditional binomial model.

While Liu and Hong (2015)'s study indicates the promise of fitting the tank car accident data by accounting for the correlations of tank car releases within the same train accident, there are still a number of unexplored questions, including at least the following:

- Would different correlation structures have different fits of the empirical tank car safety data?
- Does a particular model always have a better performance than another model, or is the model performance is dependent on the specific dataset?
- How do we measure the severity of a railroad tank car release incident? Do we use the conditional mean value or quantile statistics? How would these statistics vary in different statistical models?

This paper is intended to establish a new methodological framework for analyzing tank car releases based on historical railroad tank car accident data. In particular, we consider three alternative correlated binomial models, including Beta Binomial (BB) model, Family History (FH) model and Increment Risk (IR) model, respectively. Two independent sample datasets are used to validate and compare the performance of these models, versus the conventional binomial model. Finally, based on the model output, we analyze the mean value and quantile statistics of the probabilistic distribution of the number of tank cars releasing per train accident.

3. Statistical methodology

Derailment is a common type of freight-train accident in the United States (Liu et al., 2012; Liu, 2016). Therefore, this paper focuses on derailment-caused tank car releases. Let D_i denote the release of the i^{th} derailed tank car in a train derailment ($D_i = 1$ if this car releases and 0 otherwise). Let P_i denote its release probability (also called Bernoulli probability). As a result, the total number of tank cars releasing (denote as Y_n) given n tank cars derailed in a freight-train derailment can be expressed as:

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

The release of a derailed tank car can be viewed as a Bernoulli variable. It can be assumed that the Bernoulli indicators D_i are dependent in such a way that the conditional probability of release in any tank car releasing depends on the total number of cars releasing prior to the particular tank car. As described in Liu and Hong (2015), this assumption seems to be reasonable given the fact that the total number of cars releasing reflects the total accident kinetic energy, which is related to tank car release probability (Liu et al., 2014).

Mathematically, the above-mentioned dependency assumption is expressed as follows:

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

For illustrative convenience, we adopt a more concise notion of tank car release dependency based on a previous statistical study from Yu and Zelterman (2002):

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

where $C_n(s)$ denotes the conditional probability that the nth derailed tank car would release, given that there are s tank cars releasing prior to it. We also define $C_1 = C_1(0) = P(D_1 = 1)$. Let $P_n(s)$ ($n \geq 1$) denotes the probability of releasing s tank cars out of n derailed tank cars, that is

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

Using the Law of Total Probability (LTP), we can derive P_n using the following recursive algorithm:

$$P_n(s) = C_n(s-1)P_{n-1}(s-1) + [1 - C_n(s)]P_{n-1}(s) \tag{5}$$

Eq. (5) provides a recursive algorithm to calculate the probability