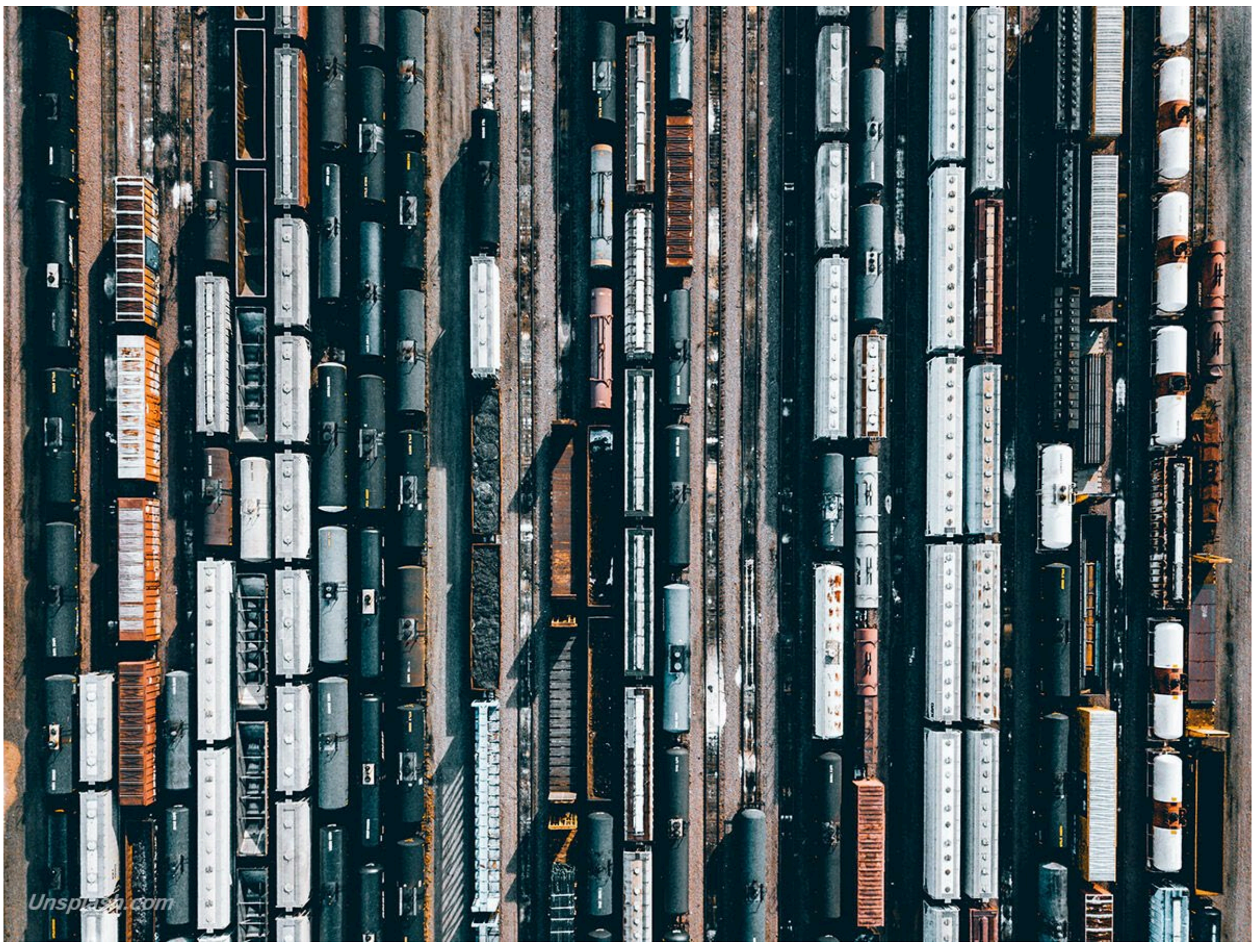
U.S. Department of Transportation

Federal Railroad Administration

Office of Research, Development and Technology
Washington, DC 20590

Development of a Bayesian Network Based Accident Model for Hazmat Unit Trains



DOT/FRA/ORD-23/37

Final Report | September 2023

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
May 2022	Technical Report	*August 2019 – May 2022*

4. TITLE AND SUBTITLE

Development of a Bayesian Network-Based Accident Model for Hazardous Material Trains

5a. CONTRACT NUMBER
693JJ619C000016

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Peter Johnson – ORCID: 0000-0002-2586-6009
Mahesh Raju Bailakanavar – ORCID: 0000-0002-8537-1673

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Thornton Tomasetti
120 Broadway 15th Floor
New York
NY 10271

8. PERFORMING ORGANIZATION REPORT NUMBER
NN19210-R01

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Department of Transportation
Federal Railroad Administration
Office of Railroad Policy and Development
Office of Research, Development, and Technology
Washington, DC 20590

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
DOT/FRA/ORD-23/37

12. DISTRIBUTION/AVAILABILITY STATEMENT
This document is available to the public through the FRA website.

13. SUPPLEMENTARY NOTES
COR: Francisco González, III

14. ABSTRACT

The objective of this research project was to develop a predictive risk model for the release of hazardous material transported in unit trains using data science techniques for available rail accident and traffic data. This work builds upon previous research (Bing, et al., 2015) which examined the causal sequence of events which can lead to a rail accident and used historical accident record and rail traffic data to define conditional probabilities of occurrence and thereby predict the risk of a hazmat release. Three different Bayesian Networks (BNs) were implemented to study the causal relationships between weather, track, and train related risk factors and the primary causes leading to railroad accidents. The BN-based accident model has demonstrated potential to accurately predict the accident cause, given information about the train and track.

15. SUBJECT TERMS: Risk analysis, unit trains, hazmat

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Francisco González, III
				48	**19b. TELEPHONE NUMBER** *(Include area code)* 202-493-6076

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

METRIC/ENGLISH CONVERSION FACTORS

ENGLISH TO METRIC

LENGTH (APPROXIMATE)

1 inch (in)	=	2.5 centimeters (cm)
1 foot (ft)	=	30 centimeters (cm)
1 yard (yd)	=	0.9 meter (m)
1 mile (mi)	=	1.6 kilometers (km)

AREA (APPROXIMATE)

1 square inch (sq in, in^2)	=	6.5 square centimeters (cm^2)
1 square foot (sq ft, ft^2)	=	0.09 square meter (m^2)
1 square yard (sq yd, yd^2)	=	0.8 square meter (m^2)
1 square mile (sq mi, mi^2)	=	2.6 square kilometers (km^2)
1 acre = 0.4 hectare (he)	=	4,000 square meters (m^2)

MASS - WEIGHT (APPROXIMATE)

1 ounce (oz)	=	28 grams (gm)
1 pound (lb)	=	0.45 kilogram (kg)
1 short ton = 2,000 pounds (lb)	=	0.9 tonne (t)

VOLUME (APPROXIMATE)

1 teaspoon (tsp)	=	5 milliliters (ml)
1 tablespoon (tbsp)	=	15 milliliters (ml)
1 fluid ounce (fl oz)	=	30 milliliters (ml)
1 cup (c)	=	0.24 liter (l)
1 pint (pt)	=	0.47 liter (l)
1 quart (qt)	=	0.96 liter (l)
1 gallon (gal)	=	3.8 liters (l)
1 cubic foot (cu ft, ft^3)	=	0.03 cubic meter (m^3)
1 cubic yard (cu yd, yd^3)	=	0.76 cubic meter (m^3)

TEMPERATURE (EXACT)

[(x-32)(5/9)] °F = y °C

METRIC TO ENGLISH

LENGTH (APPROXIMATE)

1 millimeter (mm)	=	0.04 inch (in)
1 centimeter (cm)	=	0.4 inch (in)
1 meter (m)	=	3.3 feet (ft)
1 meter (m)	=	1.1 yards (yd)
1 kilometer (km)	=	0.6 mile (mi)

AREA (APPROXIMATE)

1 square centimeter (cm^2)	=	0.16 square inch (sq in, in^2)
1 square meter (m^2)	=	1.2 square yards (sq yd, yd^2)
1 square kilometer (km^2)	=	0.4 square mile (sq mi, mi^2)
10,000 square meters (m^2)	=	1 hectare (ha) = 2.5 acres

MASS - WEIGHT (APPROXIMATE)

1 gram (gm)	=	0.036 ounce (oz)
1 kilogram (kg)	=	2.2 pounds (lb)
1 tonne (t)	=	1,000 kilograms (kg)
	=	1.1 short tons

VOLUME (APPROXIMATE)

1 milliliter (ml)	=	0.03 fluid ounce (fl oz)
1 liter (l)	=	2.1 pints (pt)
1 liter (l)	=	1.06 quarts (qt)
1 liter (l)	=	0.26 gallon (gal)
1 cubic meter (m^3)	=	36 cubic feet (cu ft, ft^3)
1 cubic meter (m^3)	=	1.3 cubic yards (cu yd, yd^3)

TEMPERATURE (EXACT)

[(9/5) y + 32] °C = x °F

QUICK INCH - CENTIMETER LENGTH CONVERSION

Inches	0	1	2	3	4	5							
Centimeters	0	1	2	3	4	5	6	7	8	9	10	11	12	13

QUICK FAHRENHEIT - CELSIUS TEMPERATURE CONVERSIO

°F	-40°	-22°	-4°	14°	32°	50°	68°	86°	104°	122°	140°	158°	176°	194°	212°
°C	-40°	-30°	-20°	-10°	0°	10°	20°	30°	40°	50°	60°	70°	80°	90°	100°

For more exact and or other conversion factors, see NIST Miscellaneous Publication 286, Units of Weights and Measures. Price $2.50 SD Catalog No. C13 10286

Updated 6/17/98

Contents

Illustrations

Tables

Executive Summary

The Federal Railroad Administration (FRA) contracted with Thornton Tomasetti to develop a predictive risk model for the release of hazardous material (hazmat) transported in unit trains using data science techniques and available rail accident and traffic data. The research was coordinated by Thornton Tomasetti from their New York offices between August 2019 and May 2022. This work builds upon previous research (Bing, et al., 2015) which examined the causal sequence of events that can lead to a rail accident. Researchers used historical accident record and rail traffic data to define conditional probabilities of occurrence and predict the risk of a hazmat release.

The research team reviewed accident and traffic data sources to inform the development of a risk model for hazmat transportation by rail. Due to the limitations of the available traffic data, the team decided to focus research on building a Bayesian-based accident model using data from FRA's Railroad Accident/Incident Reporting System (RAIRS).

The most detailed traffic data, the Surface Transportation Board (STB) Confidential Carload Waybill Samples (CCWS), provides information about car and commodity movements, but does not provide information about the train in which the car was transported, limiting its use in the research. Other traffic data reviewed, such as Class I railroad R-1 reports (Surface Transportation Board, 2021a) and the Association of American Railroads (AAR)'s Ten-Year Trends (Association of American Railroads, 2020), provide annual statistics on freight transport, including unit train miles, but do not distinguish hazmat unit trains from non-hazmat unit trains.

Consequently, the research team was unable to calculate risk, even in simple terms of accident likelihood per unit distance travelled by hazmat unit train versus hazmat manifest train. Therefore, researchers focused on building a predictive model based solely on the RAIRS accident data. The objective of the model was to predict the cause of an accident from information about the accident, based on trends and associations derived from 10 years of historical accident records. The trends and associations derived from the accident data may then indicate differences in unit train accidents compared to manifest trains and provide information about the relative risk of hazmat transportation.

The research team implemented three different Bayesian Networks (BNs) to study the causal relationships between weather, track, and train related risk factors and the primary causes leading to railroad accidents. These primary causes were selected based on the risk/consequences they posed and were then categorized in groups developed by Bing et al. (Bing et al., 2015). The team used a random forest algorithm to select the risk factors in continuous data form and then convert them into discrete categories by studying their distributions. Researchers demonstrated the capabilities of the BN-based accident risk model (e.g., predictive analytics, sensitivity analysis, and estimation of marginal and joint probability distributions) by implementing of the three separate BNs for weather, track, and train related factors. The final network integrated the train and track networks into one single network and included a variable for train type (i.e., unit train). The team identified hazmat unit trains in the RAIRS database based on the number of hazmat cars ($\geq$70) and the number of buffer cars ($\leq$5).

After reviewing the model predictions, the team found that train type did not have significant influence on the predicted accident cause, likely due to the limited scope of the current model. This model is built upon a subset of data comprising hazmat freight trains travelling on main track where a derailment occurred and the cause was one of four specific cause groups: Wide Gauge (03T), Track Geometry (04T), Buckled Track (05T), Broken Rail, or Welds (08T). This limits the number of unit train accidents in the data set, which consequently limits the ability to identify differences in accident properties.

Researchers used accident data from the year 2019 and 2020 to validate the integrated BN and applied data filters used for developing the Sort-C1 on this data; in total, 24 accidents were selected. An important aspect of the validation is that the BN had not seen this data previously. The risk factors from the 24 accidents were provided as inputs to the BN and the network was tasked to make blind predictions on the cause for these accidents. The integrated network predicted with 100 percent accuracy for accident cause groups 03T and 08T and 79 percent accuracy for 04T. The lower accuracy for cause group 05T is attributed to the small amount of training data in Sort-C1. However, the overall accuracy of the network predictions was about 79 percent, which appears promising. Increasing the training data pool (i.e., Sort C1) is expected to improve this accuracy of the network.

Improvements in accident and traffic data collection are required to aid future research on the risk of hazmat transportation by unit train. The team recommends a new field be added to the FRA RAIRS database, indicating whether the train is a unit train. However, since researchers observed significant variation in how a unit train is defined, such a change would require adoption of a standard definition of a unit train.

It may not be feasible to collect the level of information required to fully incorporate traffic data into the architecture of the Bayesian-based accident model, even in a geographically limited model. This would require that all successful train journeys over a given period be added into the database alongside those which ended in an accident. However, for more traditional statistical methods, improvements could be made to identifying hazmat unit train traffic. Schedule 755 of the R-1 reports (i.e., Railroad Operating Statistics) submitted annually to the STB (Surface Transportation Board, 2021a) includes data on total train-miles and car-miles by train type, including unit trains. Therefore, the team recommends that the collection of railroad operating statistics be expanded to provide total train-miles and car-miles by train type transporting hazmat.

The BN-based accident model has demonstrated potential to accurately predict accident cause when provided with information about the train and track. The team recommends further development of the model's capabilities by the incorporation of additional cause groups (e.g., mechanical, electrical, and human causes) and the inclusion of track maintenance inspection and repair data, which could enhance causal relationship learning.

1. Introduction

The Federal Railroad Administration (FRA) contracted with Thornton Tomasetti to develop a predictive risk model for the release of hazardous material (hazmat) transported in unit trains using data science techniques and available rail accident and traffic data. The research was coordinated by Thornton Tomasetti from their New York offices between August 2019 and May 2022.

1.1 Background

Freight railroads continually strive to increase their productivity by transporting greater amounts of goods in the shortest amount of time. Consequently, average axle loads and train speeds have tended to increase over time. Freight railroads also boost their productivity by operating unit trains. A *unit train* is a train transporting the same commodity from the same origin to the same destination. Railroads operate unit trains to increase efficiency and productivity by reducing costs, employing bulk loading, improving asset utilization, and reducing transit time. The use of unit trains began to be prevalent in the 1950s to transport large shipments of coal. By the end of the 1960s, approximately 90 percent of all coal traffic on U.S. railroads was transported via unit trains (Starr, 1976). Railroads now operate unit trains for bulk shipments of commodities such as iron ore; corn, wheat, and grain; sand and gravel; garbage and liquid sludge; automobiles; steel billets; citrus fruits and vegetables; and hazardous materials. However, the risk of using unit trains to transport hazmat has come under scrutiny after several recent railroad hazmat accidents. Figure 1 shows a timeline of railroad accidents in the U.S. involving the release of crude oil and ethanol between 2009 and 2018.[1]

Table 1 lists the accident date, accident location, commodity released, and whether a unit train was involved. The table links to National Transportation Safety Board (NTSB) accident reports.

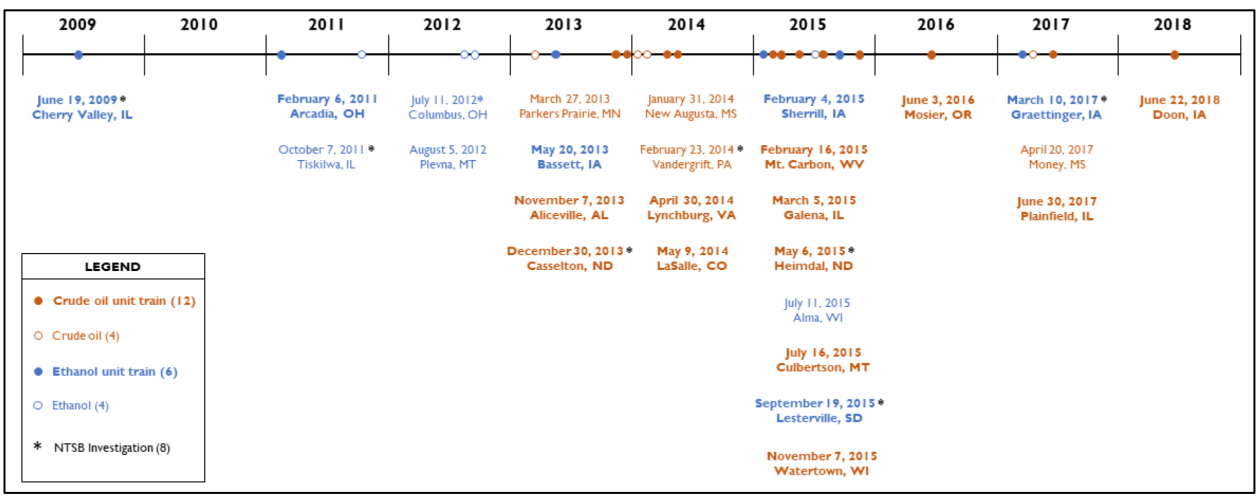


Figure 1. Timeline of Crude Oil and Ethanol Railroad Accidents in the U.S. 2009-2018

[1] This ten-year period was the targeted timeframe for the current project. It should be noted that several accidents involving the release of crude oil occurred in Canada during this timeframe, including one of the deadliest in Canadian railroad history (Lac-Mégantic, Quebec on July 6, 2013).

Table 1. List of Crude Oil and Ethanol Railroad Accidents in the U.S., 2009-2018

Date	Location	Hazmat	Unit Train (Y/N)	NTSB Report No.
June 19, 2009	Cherry Valley, IL	Ethanol	Y	RAR1201
February 6, 2011	Arcadia, OH	Ethanol	Y	-
October 7, 2011	Tiskilwa, IL	Ethanol	N	-
July 11, 2012	Columbus, OH	Ethanol	N	RAB1302
August 5, 2012	Plevna, MT	Ethanol	N	
March 27, 2013	Parkers Prairie, MN	Crude Oil	N	RAB1408
May 20, 2013	Bassett, IA	Ethanol	Y	-
November 7, 2013	Aliceville, AL	Crude Oil	Y	-
December 30, 2013	Casselton, ND	Crude Oil	Y	-
January 31, 2014	New Augusta, MS	Crude Oil	N	-
February 23, 2014	Vandergrift, PA	Crude Oil	N	-
April 30, 2014	Lynchburg, VA	Crude Oil	Y	RAB1701
May 9, 2014	LaSalle, CO	Crude Oil	Y	-
February 4, 2015	Sherrill, IA	Ethanol	Y	-
February 16, 2015	Mount Carbon, WV	Crude Oil	Y	-
March 5, 2015	Galena, IL	Crude Oil	Y	RAB1601
May 6, 2015	Heimdal, ND	Crude Oil	Y	-
July 11, 2015	Alma, WI	Ethanol	N	-
July 16, 2015	Culbertson, MT	Crude Oil	Y	-
September 19, 2015	Lesterville, SD	Ethanol	Y	-
November 7, 2015	Watertown, WI	Crude Oil	Y	-
June 3, 2016	Mosier, OR	Crude Oil	Y	RAB1712
March 10, 2017	Graettinger, IA	Ethanol	Y	-
April 20, 2017	Money, MS	Crude Oil	N	-
June 30, 2017	Plainfield, IL	Crude Oil	Y	-
June 22, 2018	Doon, IA	Crude Oil	Y	RAB1707

Bing et al. (2015) examined Railroad Accident/Incident Reporting System (RAIRS) data for a five-year period between 2004 and 2008 (inclusive) to study the causes of freight train accidents and derailed freight cars. The use of unit trains to transport hazmat was not included in this study but was recommended for future research.

Since this time, research has been conducted to examine the risk of operating unit trains (Liu, 2017) (Li et al., 2018). Li et al. (2018) studied the 15-year period between 2001 and 2015. However, the focus of this study was to characterize the common modes of derailment for empty and fully loaded unit tank cars.

1.2 Objectives

The objective of this research project was to develop a predictive risk model for the release of hazardous material transported in unit trains using data science techniques for available rail accident and traffic data.

1.3 Overall Approach

This work builds upon previous research (Bing et al., 2015) which examined the causal sequence of events which can lead to a rail accident. Researchers used historical accident record and rail traffic data to define conditional probabilities of occurrence and thereby predict the risk of a hazmat release. The team used available data collected by FRA, railroads, and track inspection service providers on accidents, rail traffic, and track condition. Researchers used artificial neural networks and data science approaches to identify predictable patterns, with a particular focus on unit train accidents. These patterns and the process of their identification provided a foundation for accident prediction and were incorporated into a predictive model for hazmat release.

1.4 Scope

The scope of the research was limited to transportation of hazardous goods via freight train in the United States, excluding all other train traffic, analyzed using artificial neural networks and data science approaches. The team used accident data from FRA's RAIRS and obtained detailed traffic data from the Surface Transportation Bureau (STB)’s Confidential Carload Waybill Samples (CCWS) (Surface Transportation Board, 2021b), summary data from STB R-1 annual reports (Surface Transportation Board, 2021a), and the Association of American Railroads (AAR) Ten-Year Trends report (Association of American Railroads, 2020).

1.5 Organization of the Report

Section 2 describes the data reviewed for the research. Section 3 describes the development and verification of the Bayesian-based accident model. Section 4 presents conclusions and recommendations.

2. Data Review

The objective of the research was to understand the relative risk of transporting hazmat by unit train versus manifest train. In the context of train accidents, risk is expressed as a likelihood of an accident per unit distance travelled by a train, typically expressed in miles. This requires information on accidents (i.e., the numerator) and rail traffic (i.e., the denominator). A method is required to identify hazmat unit trains and hazmat manifest trains in both accident data and rail traffic data.

The following section provides an overview of the data reviewed for this research and how it was used to meet the central objective of the study.

2.1 RAIRS Accident Database

2.1.1 Overview

FRA's RAIRS includes a database of accident records comprising 200,000 accidents between 1975 and 2018. For the development of a predicative model, a date range must be selected, and recent data was considered preferable since it is representative of current standards.

The team performed an initial review of the accident data and found a significant variation in accident number and cause group since 1975, showing an overall trend of a decrease in the number of accidents per year (Figure 2). In the years 2009-2018 (inclusive), the number and causes of accidents has been relatively consistent; therefore, this date range was chosen for the model.

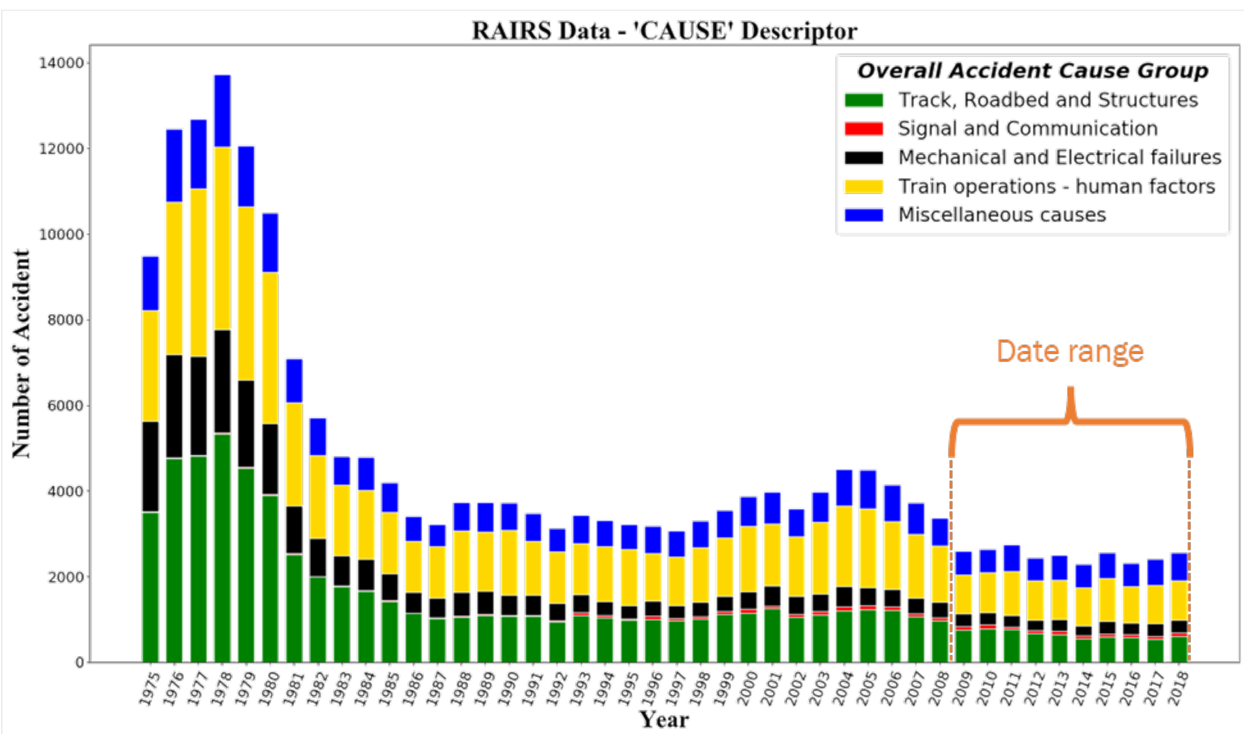


Figure 2. Trend of Accident Records from 1975 to 2018 Showing Primary Accident Cause Group

2.1.2 Accident Data Filtering

The RAIRS database was sorted into several subsets for use in the analysis, filtered down to the group containing those in which hazmat unit trains were found.

- Sort-A = All accident entries from 2009-2018
- Sort-B = Sort-A and the train must contain a hazmat car (i.e., CARS > 0)
- Sort-C = Sort-B and must comprise a freight train on main track (i.e., TYPEQ = 1 and TYPTRK = 1)
- Sort-D = Sort-C and a hazmat car released product (i.e., CARSHZD > 0)

The number of accident entries in RAIRS for each sorting group is listed in Table 2, where each entry refers to a train involved in an accident. For multiple trains involved in a single accident, there is an entry for each train.

For comparison of unit trains versus manifest trains carrying hazmat, Sort-C is the most relevant. Sort-C filters out trains that are not carrying hazmat and accidents on track types that are not typically used by unit trains of any type.

Table 2. Entries in RAIRS Database for the Four Sorting Groups, 2009-2018

Year	Sort-A	Sort-B	Sort-C	Sort-D
2009	2,597	609	180	12
2010	2,641	621	192	13
2011	2,748	662	216	8
2012	2,429	599	167	22
2013	2,497	603	202	14
2014	2,280	539	179	12
2015	2,553	553	154	14
2016	2,309	477	159	7
2017	2,403	526	184	9
2018	2,563	522	183	8
Total	25,020	5,711	1,816	119

It should be noted that the statistical methods proposed for this research are best suited to datasets several orders of magnitude larger than the Sort C group (i.e., the 2009-2018 date range, or 1,816). While this limitation was noted, the team decided to proceed with this date range for initial research and then once the model was completed, conduct a model validation exercise to show the accuracy of the model and determine whether there was justification to increase the date range.

2.1.3 Identification of Unit Trains

The RAIRS data fields do not include an indicator for accidents involving a hazmat unit train, nor does it include fields from which one can directly infer whether the train is carrying a single commodity or whether all cars are transported from a single origin to a single destination. Furthermore, while a general definition of a unit train is well understood, several different hazmat unit train definitions were identified. A combination of these were ultimately used to arrive at working definition of a hazmat unit train and this was applied to the RAIRS Sort C data.

The initial approach taken to identify hazmat unit trains in the RAIRS database included review of each of the accident narratives in the Sort D dataset (i.e., 119 entries) to identify those where hazmat unit trains were recorded as involved in the accident event, relying primarily on information provided in NTSB accident reports and news reports. This review identified a total of 23 hazmat unit trains in the Sort D dataset. A review of the trains' common characteristics showed that they were typically comprised of many cars, ranging between 61 and 116 cars in total. Also, all 23 trains were almost entirely comprised of loaded hazmat freight cars, and the number of buffer cars (i.e., cars which were recorded as not carrying hazmat) ranged between 1 and 3. Consequently, review indicated that both train length and number of buffer cars could be used to identify unit trains in the RAIRS database.

Researchers ultimately used the unit train definition provided in Title 49 (Transportation) of the Code of Federal Regulations (CFR)[2] to identify unit trains. This describes a High-Hazard Flammable Unit Train (HHFUT) as *"a single train transporting **70 or more** loaded tank cars containing Class 3 flammable liquid."* Using this description to analyze the Sort D dataset, "≥70 cars" was determined to be an appropriate ruleset for identifying a unit train in the RAIRS database.

NTSB safety recommendation report R-17-01 (National Transportation Safety Board, 2020) recommended *"positioning placarded railcars in a train and require that all trains **have a minimum of five non-placarded cars** between any locomotive or occupied equipment and the nearest placarded car transporting hazardous materials, regardless of train length and consist."* Therefore, trains ≥70 cars in length and containing ≤5 buffer cars were flagged as unit trains in the RAIRS database.

In summary, for the Sort C RAIRS dataset, which comprises freight trains on main track transporting hazmat, the following filters were applied to the RAIRS database to identify unit trains:

- (Total number of cars) – (Loaded hazmat cars + empty hazmat cars) <= 5
- (Total number of cars) >= 70

Further information on train type was obtained for Class I railroads and their subsidiaries. Of the 1,816 records in the Sort C dataset, 1,594 (88 percent) included trains operated by Class I railroads or their subsidiaries. These definitions were not used as a general ruleset, serving instead as a verification of the adopted ruleset. A summary of unit train definitions obtained for Class I railroads is provided in Appendix A.

2.2 Traffic Data

2.2.1 STB Confidential Carload Waybill Samples (CCWS)

The CCWS is collated by STB. It is a stratified sample of carload waybills for all U.S. rail traffic submitted by those rail carriers terminating 4,500 or more revenue carloads annually (Surface

[2] Code of Federal Regulations (CFR), Title 49 (Transportation), Subtitle B, Chapter I, Subchapter C, Part 171, Subpart A, Section 171.8 – Definitions and abbreviations

Transportation Board, 2021b). Upon request, STB provided confidential carload waybill samples from 2009-2018 (inclusive), in line with the range of accident data considered, to assist with the research.

Each line of the database describes the movement of a carload, including details on origin, destination, material transported, and date transported. However, the data is limited in that it does not provide any information on the train with which any individual car was transported. This information cannot be determined by using common dates, origins, and destinations to try to reconstruct a theoretical train; because the dates are from accounting data and not operational data, the carload may not have moved on the same day. Therefore, there is no means of identifying historical traffic by train type using the CCWS data, and no means of discerning relative risk for transporting hazmat by unit train or manifest train.

2.2.2 Summary Reports

All Class I railroads are required to submit annual R-1 reports to the STB, which are made available to the public via the STB website (Surface Transportation Board, 2021a). These reports provide data on total train miles and car miles for unit trains and manifest trains (Subsection "*Schedule 755 – Railroad Operating Statistics"*). Similarly, AAR's Ten-Year Trends (Association of American Railroads, 2020) provides annual statistical summaries of rail transportation derived from various STB annual reports. However, neither of these sources provide a breakdown of hazmat transportation by unit train or manifest train.

2.3 Data Review Discussion

The team reviewed both accident data and traffic data sources to assist in the development of a risk model for hazmat transportation by rail. However, due to the limitations of the available traffic data, researchers decided to build a model using only RAIRS accident data.

One of the constraints of the research was determining how to apply a consistent analytical method to accident and traffic datasets, which necessarily contain different information. For example, each row of the RAIRS database contains information about a train which was involved in an accident, with each column describing some feature of the train or the accident. Using this information, the team proposed a model which could predict the cause of an accident given information about the accident, based on trends and associations derived from 10 years of historical accident records. The trends and associations derived from the accident data may then indicate differences in unit train accidents compared to manifest trains when analyzing the relative risk of hazmat transportation.

Using only the RAIRS database, this model would not be able to calculate risk (i.e., the likelihood of an accident per unit distance travelled by unit train or manifest train). This requires information on overall traffic to contextualize the number of accidents for hazmat unit trains and hazmat manifest trains. To include traffic data, the RAIRS data would have to be supplemented with additional entries as rows in the dataset, where each entry describes a successful journey with equivalent data regarding the train and journey (e.g., train makeup, origin, destination, etc.) in each column. The STB CCWS data is the most detailed traffic information available, but it is limited in that it represents only a sample of journeys and does not provide information about the train in which the car travelled in. The former issue could be addressed by scaling the number of entries in the augmented RAIRS database based on the sampling rate used in the STB CCWS. However, the latter cannot be surmounted, as each entry in the augmented RAIRS database would represent a train, and there is no way to link the cars recorded in the STB CCWS to a train or determine whether it was transported by unit train or a manifest train.

Other traffic data sources covering 2009-2018, such as R-1 reports and AAR's Ten-Year Trends (Association of American Railroads, 2020), provide only annual statistical summaries of Class I railroad train traffic and could not be incorporated into the RAIRS-based model due to a lack of detail on individual train movements. These data sources provide annual traffic statistics in terms of train miles and car miles for all freight train traffic and unit train traffic, so the team investigated whether a simple statistical risk calculation could be conducted to determine the accident frequency per mile travelled by the average hazmat unit train and compare this with the accident frequency for the average hazmat manifest train. However, neither the R-1 reports nor AAR's Ten-Year Trends (Association of American Railroads, 2020) provide information on whether the unit trains are hazmat unit trains.

In summary, after a review of available data, the team concluded that the central objective of the research could not be met because it required more detailed information, particularly on hazmat unit train traffic. Therefore, the primary focus of the research became the development of a Bayesian-based accident model, based on RAIRS data between 2009-2018, as described in the following section.

3. Bayesian Network Based Rail Accident Model

3.1 Bayesian Networks

Bayesian Networks (BNs) are directed acyclic graphs (DAGs) consisting of nodes representing random variables and arrows that correspond to the probabilistic cause-effect (i.e., parent-child) relationship between the random variables. BNs are versatile; the lines between boxes are agnostic to the method of characterization and robust enough to handle small and incomplete data sets. Statistical or computational methods are used to estimate the conditional dependencies between the random variables. BNs are powerful tools for knowledge representation, reasoning, and modeling the causal relationships of a given effect. BNs allow for integrating historical data with a subject matter expert (SME)'s knowledge to visually show the probabilistic relationships among a set of random variables.

BNs have been used extensively for casual relationship modeling, risk assessment, decision making, and uncertainty quantification across industry sectors. They have been used in medical diagnostics (Heckerman et al., 1995), environmental modeling (Uusitalo, 2007), civil engineering and construction management risk analysis (Fan & Yu, 2004; Luu et al., 2009) (Zhang et al., 2014), structural health monitoring (Zhang et al., 2016), and accident risk analysis (Cheng et al., 2010; Hänninen, 2014; Camino López et al., 2008; Martín et al., 2009).

In this research, the BN model was implemented to establish the probabilistic cause and effect relationship between railroad accidents and various contributing factors, and to model any interdependencies that exist between the causal factors themselves. It is important to note that the accident cause descriptor used in the RAIRS database (i.e., CAUSE and CAUSE2 columns) is the end effect and the rest of the 143 parameters recorded post-accident are considered causal factors. There are five key stages to the implementation of the BN-based accident model as set out below:

1. RAIRS Accident Data Sorting
2. Railroad Accident Causes and Risk Factors Identification
3. Risk Factors – Continuous to Categorical Data Conversion
4. BN-based Railroad Accident Risk Model Implementation
5. Railroad Accident Risk Model Validation

3.2 RAIRS Accident Data Sorting

3.2.1 Accident Statistics

This section discusses the accident statistics derived from the Sort-C and Sort-D data and the trends observed.

Figure 3 shows the number of train accidents and the number of hazmat cars releasing hazmat for each year between 2009 and 2018. Over this period, a total of 388 cars released hazmat in 119 freight train accidents. While the total number of hazmat cars releasing each year fluctuated over this time span, the number of accidents per year was relatively constant except for a spike in 2012. The breakdown of the accidents by railroad type and track class for Sort-C and Sort-D is shown in Figure 4 and Figure 5, respectively. The figures show that most of the Class I railroad accidents occurred on Class-4 track, followed by Class-3 and Class-2 track.

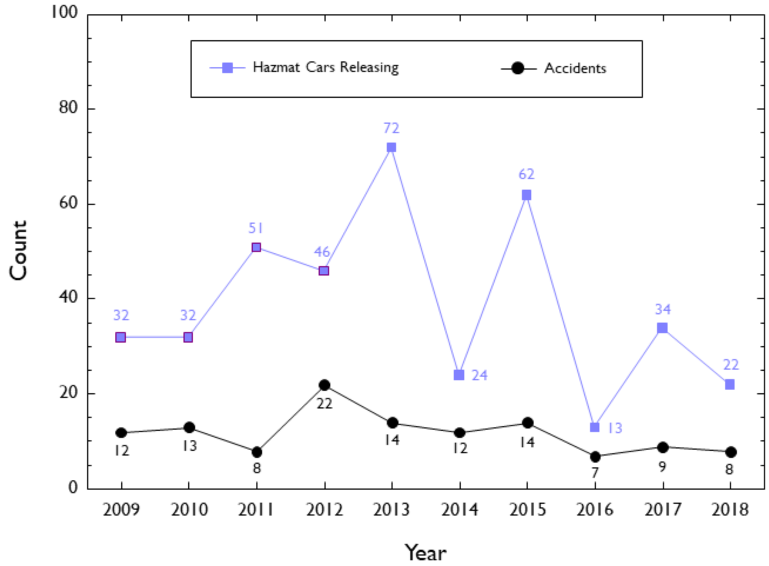


Figure 3. RAIRS Freight Train Accidents and Hazmat Cars Releasing on Main Track, 2009-2018

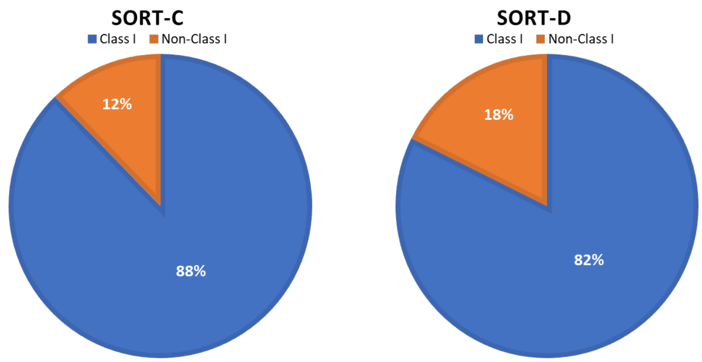


Figure 4. Distribution of Accidents by Railroad Type

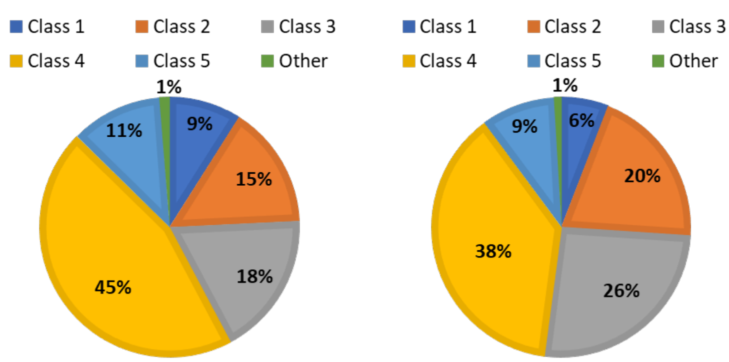


Figure 5. Distribution of Accidents by FRA Track Class

A breakdown of the Sort-C accident data by accident type is shown in Table 3. Approximately 60 percent of the accidents in Sort-C are derailment type accidents followed by highway rail crossing at 26 percent.

Table 3. Sort C Accidents Breakdown by Accident Type

Accident Type Code	Accident Type Description	Sort C Count	Sort C %
1	Derailment	1073	59%
7	Highway Rail Crossing	470	26%
13	Other	64	4%
9	Obstruction	58	3%
3	Rear End Collision	42	2%
11	Fire/Violent Eruption	39	2%
4	Side Collision	21	1%
12	Other Impacts	18	1%
5	Raking Collision	12	1%
2	Head On Collision	10	1%
6	Broken Train Collision	5	0%
8	Railroad Grade Crossing	3	0%
10	Explosive Detonation	1	0%
Total		**1816**	

3.3 Railroad Accident Causes and Risk Factors Identification

The following terminologies are used in the discussion of the implementation of the BN for rail accident risk analysis.

Risk Factors, also referred to as Causal Factors, are the parameters that are recorded in the RAIRS accident database (~146 parameters) following a railroad accident. For the BN implementation these factors are grouped into the following categories:

- Weather and climate factors
- Track factors (e.g., track density, track class)
- Rail factors (e.g., gross tons, number of loaded freight cars, total number of cars)
- Tonnage and Train Composition

Accident Causes (Effect) are the causes that lead to a railroad accident that are considered to be the end effect. In this context, a broken joint bar (cause code T214) or a roller type journal bearing failure due to overheating (cause code E53C) are the end effects that can lead to an accident.

3.3.1 Data Culling - Accident Cause Categories

The RAIRS accident database contains more than 400 unique cause codes to identify the primary reason for train accidents and derailments. These codes are categorized under five major cause groups as shown in Table 4 and Figure 6.

Table 4. Accident Cause Groups

Cause Group Description	Cause Group Code
Track, Roadbed and Structures	T
Signal and Communication	S
Train Operation – Human Factor	H
Mechanical and Electrical Failures	E
Miscellaneous Causes	M

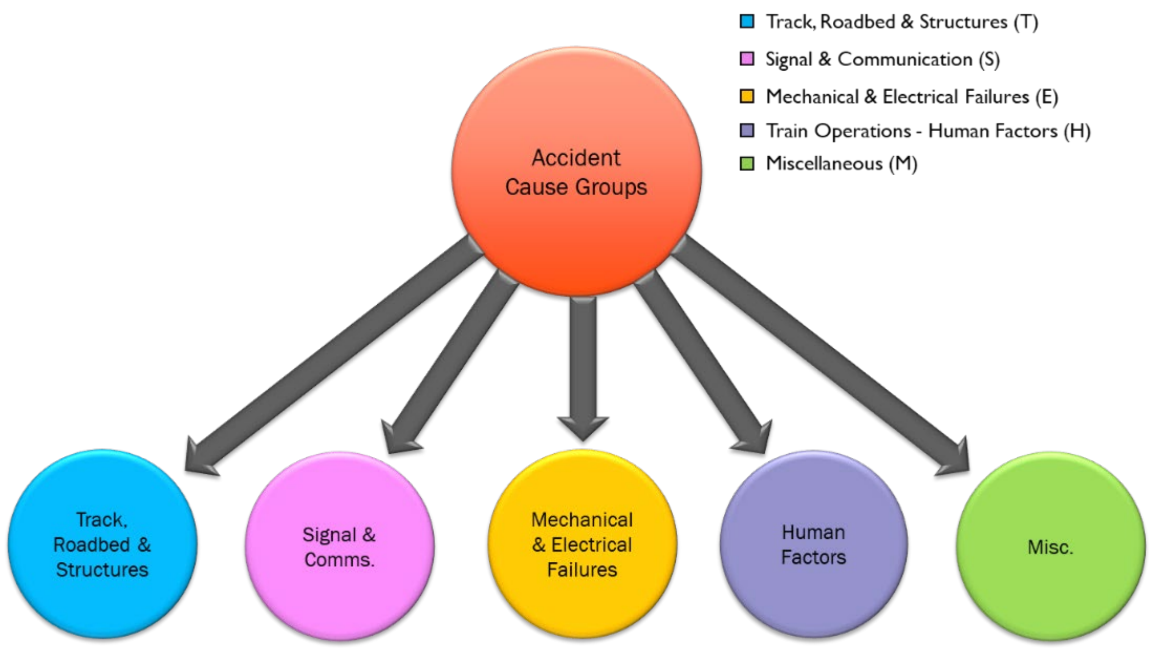


Figure 6. Illustration of the Accident Cause Groups

Table 5 shows the distribution of accidents across the cause groups over the period 2009 to 2018. The track, roadbed, and structure group is the leading cause for railroad accidents where a hazmat car released product.

Table 5. Distribution of accident causes across the four accident data sorts

Cause Group	Sort A	Sort B	Sort C	Sort D
Track, Roadbed, and Structures	6,525	1,496	416	64
Mechanical and Electrical Failures	2,776	711	408	19
Train Operation – Human Factor	9,260	2,178	294	14
Miscellaneous Causes	5,758	1,143	688	22
Signal and Communication	701	183	10	0
Total	25,020	5,711	1,816	119

The Sort C accident data pertaining to Track, Roadbed, and Structures (T), Mechanical and Electrical Failures (M), and Train Operations – Human Factor (H) was analyzed to identify cause categories with high accident counts and/or high consequence (i.e., risk). The consequence/risk is the monetary loss incurred due to a train accident or derailment. The total reportable damage (ACCDMG) data available in the RAIRS database was used to calculate the consequence/risk:

$$Category\ Risk = \frac{Category\ accident\ count}{Group\ accident\ count} * Category\ average\ accident\ loss$$

where the average loss for an accident category is given by

$$Category\ average\ accident\ loss = \frac{Category\ total\ loss}{Categoryaccident\ count}$$

The plot of accident count and consequence for each of the cause categories under T, M, and H cause groups is shown in Figure 7 and Figure 8, respectively. The plot of accident count versus consequence for the same categories is shown in Figure 9.

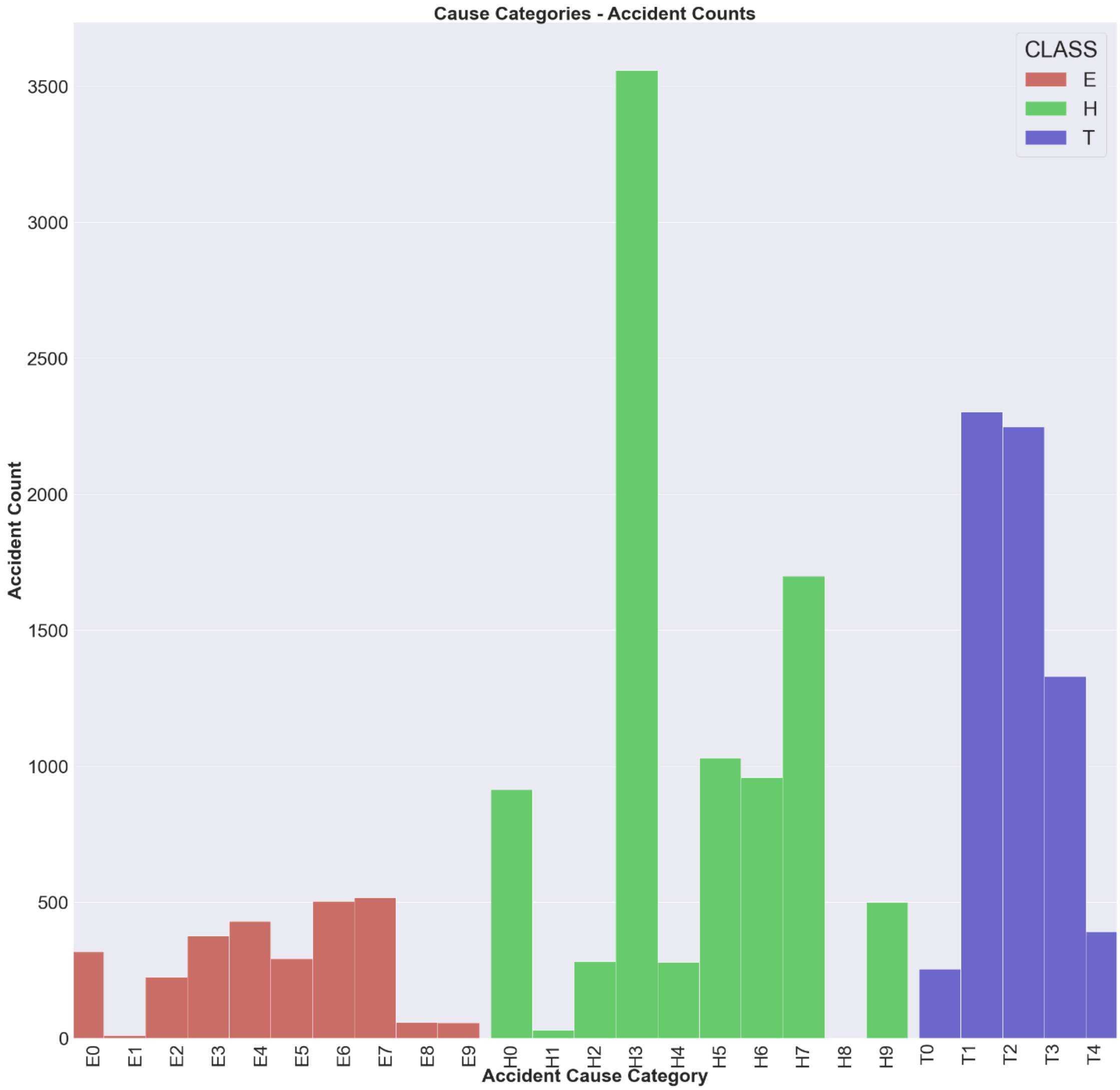


Figure 7. Accident Counts for Categories Under T, M, & H Cause Groups

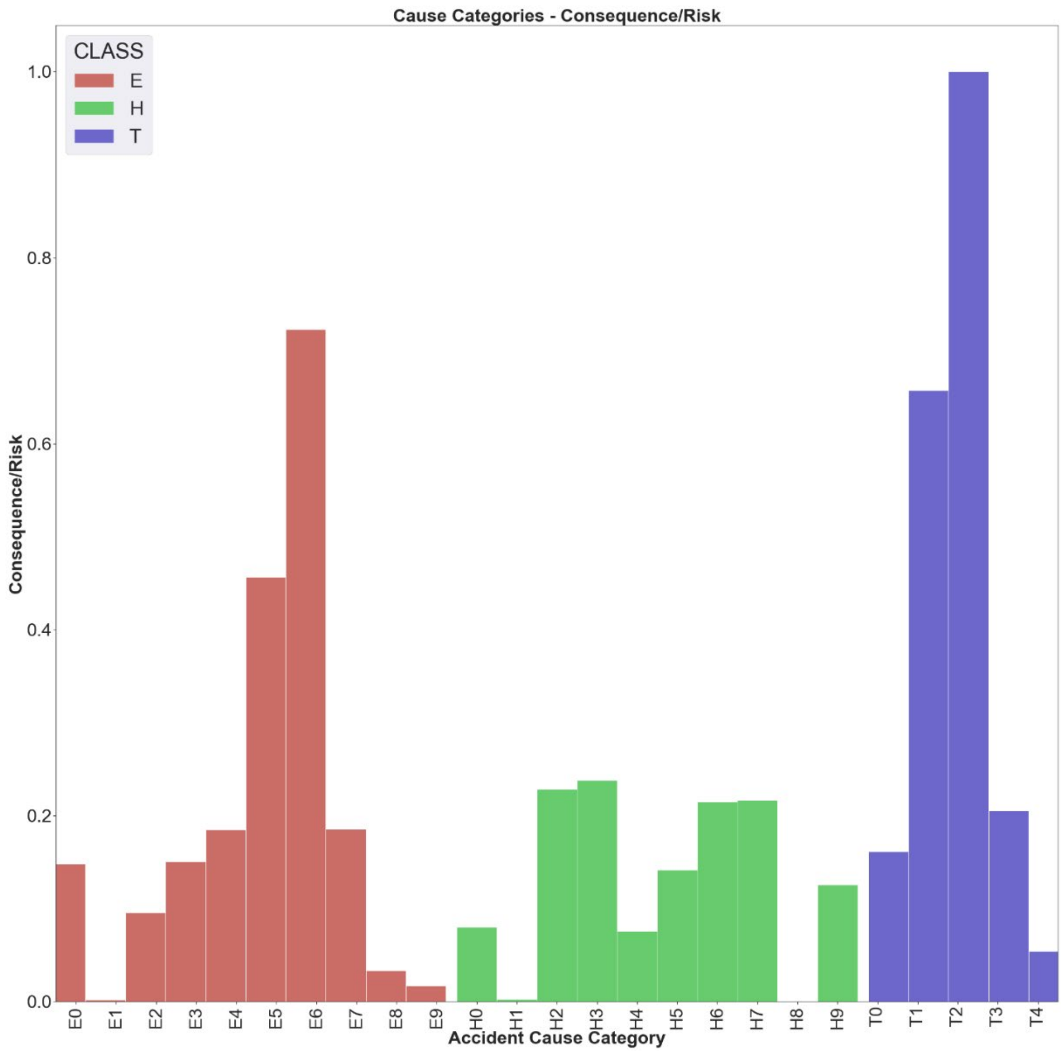


Figure 8. Consequence/Risk for Categories Under T, M, & H Cause Groups

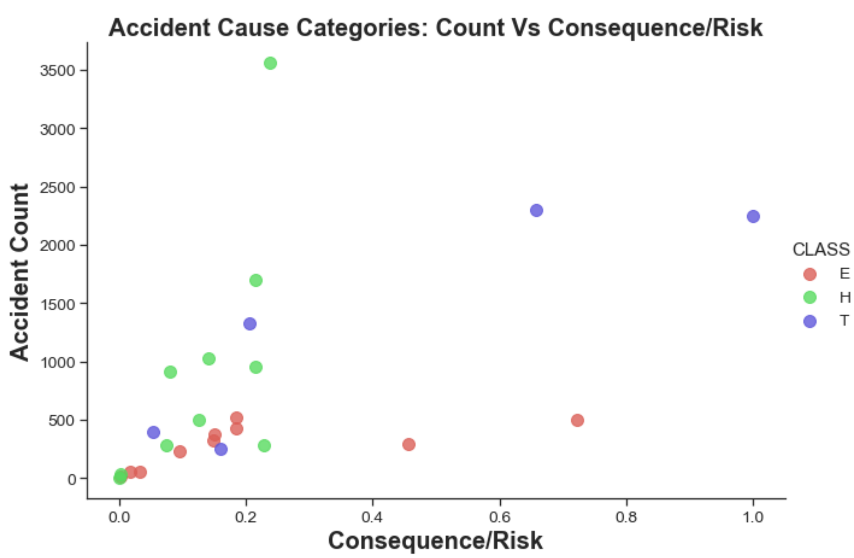


Figure 9. Accident Count vs. Consequence for Categories Under T, M, & H Cause Groups

Based on the analysis, the top-2 high-count, high-risk categories for the T, M, and H cause groups are shown in Table 6, Table 7, and Table 8, respectively. Since track, roadbed, and structure related issues pose the greatest risk/consequence, the BN-based accident risk model will only consider the causes pertaining to this group, specifically cause categories T1 and T2. In addition, the current study only considers derailments since this is the leading type of accident by counts. The implementation of BNs to study causal relationships for the other cause groups and accident types is outside the scope for this research. After the derailment accident type and track, roadbed, and structure (T) cause groups filter is applied, the total number of accidents in Sort-C is reduced to 409 accidents. For ease of reference, the reduced Sort-C database with 409 accidents is referred to as Sort-C1.

Table 6. Top-2 Categories for Track, Roadbed, and Structures (T) Cause Group

Category Description	Category Code
Track Geometry	T1
Rail, Joint Bar, and Rail Anchoring	T2

Table 7. Top-2 Categories for Mechanical and Electrical Failures (M) Cause Group

Category Description	Category Code
Axles and Journal Bearings	E5
Wheels	E6

Table 8. Top-2 Categories for Train Operations-Human Error (H) Cause Group

Category Description	Category Code
General Switching Rules	H3
Switches, Use of	H7

To simplify the implementation of the Bayesian-based accident model, the team used the grouping of accident causes by Alan Bing and others (Bing et al., 2015). The distribution of accidents related to track, roadbed, and structures in Sort C as per this grouping is presented in Figure 10. The top-four accident groups (i.e., 08T, 04T, 05T, and 03T) comprise approximately 75 percent of the accidents in Sort-C1. The number of accidents in the remainder of the groups is too small to fit the accident model and therefore removed in the implementation of the BN.

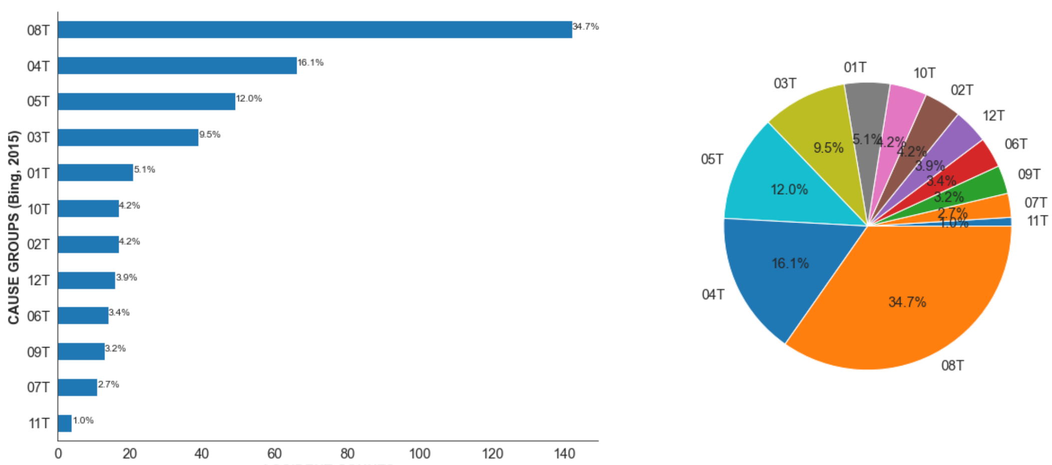


Figure 10. Distribution of Sort-C Accidents Across Cause Groups (Bing, 2015)

3.3.2 Feature Selection – Rail Operations, Weather and Climate, Tonnage, and Train Composition

The RAIRS accident database consists of approximately 146 parameters that are recorded in the event of an accident or a derailment. For data-driven risk analysis, it is important to identify the most important factors contributing to rail accidents and derailments and discard the insignificant parameters, a process called Feature Selection in Machine Learning (ML). Feature Selection helps reduce the number of input variables (i.e., reduce curse of dimensionality), reduce the computational cost, and in some cases improve the performance of the ML model. The terms *features*, *parameters*, or *factors* will be interchangeably used throughout this report to refer to the 146 recorded parameters.

The team used the Random Forest model to identify the important factors affecting the Track, Roadbed, and Structures (T), Mechanical and Electrical Failures (M), and Human Factor (H) accident causes. The concept of decision trees is central to the implementation of random forests for feature selection. A decision tree consists of a root node, intermediary nodes, and leaf nodes all connected by branches through which information flows down the tree (see Figure 11). The decision tree is built recursively by splitting the data to make use of the features. For the rail accident data, this is akin to splitting (i.e., bagging) the accident data into different accident cause categories by making use of the 146 recorded parameters. At each node the feature that best splits the data is evaluated by using metrics like Gini impurity or entropy for categorical data and residual or mean squared error (MSE) for continuous data. Splitting the data at a node results in the creation of two intermediary nodes, the first node that contains data split by the selected feature and the second node where the data is yet to be split or bagged. The process of splitting the data at the second node is then repeated. This data splitting process is recursively done until there is no data left to be split or bagged.

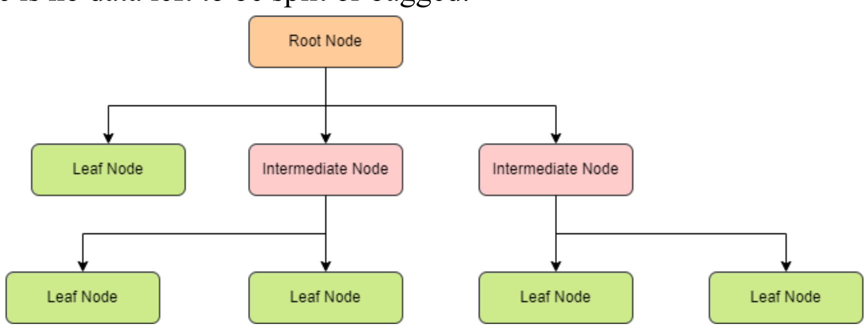


Figure 11. Illustration of a Decision Tree (Source: https://www.diagrams.net/)

The Random Forest method involves building several hundred such decision trees, each built by the random extraction of features and accident data pertaining to those features. This ensures that not every tree sees all the features or all the observations, thereby reducing overfitting. The significance of each feature based on how well it can split the data (Gini impurity or entropy for categorical data and residual or MSE for continuous data) is averaged over all the trees to determine the overall significance of the feature.

The top five parameters identified for each of the three cause groups are shown in Table 9 and Table 10. The factors affecting the Track, Roadbed, and Structures (T) and Mechanical and Electrical Failures (M) are identical and in the same order of priority. For Human Factors (H), the type of track takes precedence over the month during which the accident occurred. The remaining four factors are identical to the T & M cause groups but not in the same order.

Table 9. Top Five Factors Affecting Track, Roadbed, and Structures (T) and Mechanical and Electrical Failures (M)

Parameter Description	Parameter Code
Train speed	TRNSPD
Temperature in degrees Fahrenheit	TEMP
Number of loaded freight cars	LOADF1
Gross tonnage excluding power units	TONS
Month of incident	MONTH

Table 10. Top Five Factors Affecting Train Operations – Human Factors (H)

Parameter Description	Parameter Code
Train Speed	TRNSPD
Number of loaded freight cars	LOADF1
Type of track	TYPTRK
Temperature in degrees Fahrenheit	TEMP
Gross tonnage excluding power units	TONS

3.4 Risk Factors - Continuous to Categorical Data Conversion

The BN for rail derailment risk analysis in this study was implemented in the open-source software pomegranate[3] (Schreiber, 2018). Since pomegranate currently supports discrete data only, the continuous data fields (i.e., accident factors) in the Sort C dataset were converted to discrete variables. Histograms were used to visualize the data for each field, while grouping of the continuous data into categorical data was performed iteratively to ensure that the number of bins (i.e., categories) captured the distribution of the data. The bins were named with appropriate labels to match the distribution of the data.

Factors such as railroad company, track class, weather, and visibility are recorded as categories in the RAIRS database. Only the distribution of these variables across each of the categories is presented for these factors. For train speed, the categories were based on the speed limits for each track class. For each of the risk factors used in the implementation of the BN, their categories and the distribution of accidents across the categories is shown in Table 11 to Table 18 and Figure 12 to Figure 19.

Table 11. Train Speed (TRNSPD) Categories

Track Class	Speed Range	Train Speed Category
Class 1	Min to 10 mph	LT10
Class 2	10 to 25 mph	10TO25
Class 3	25 to 40 mph	25TO40
Class 4	40 to 60 mph	40TO60
Class 5	60 to 80 mph	60TO80

[3] pomegranate is a Python package that implements fast and flexible probabilistic models ranging from individual probability distributions to compositional models such as Bayesian networks and hidden Markov models.

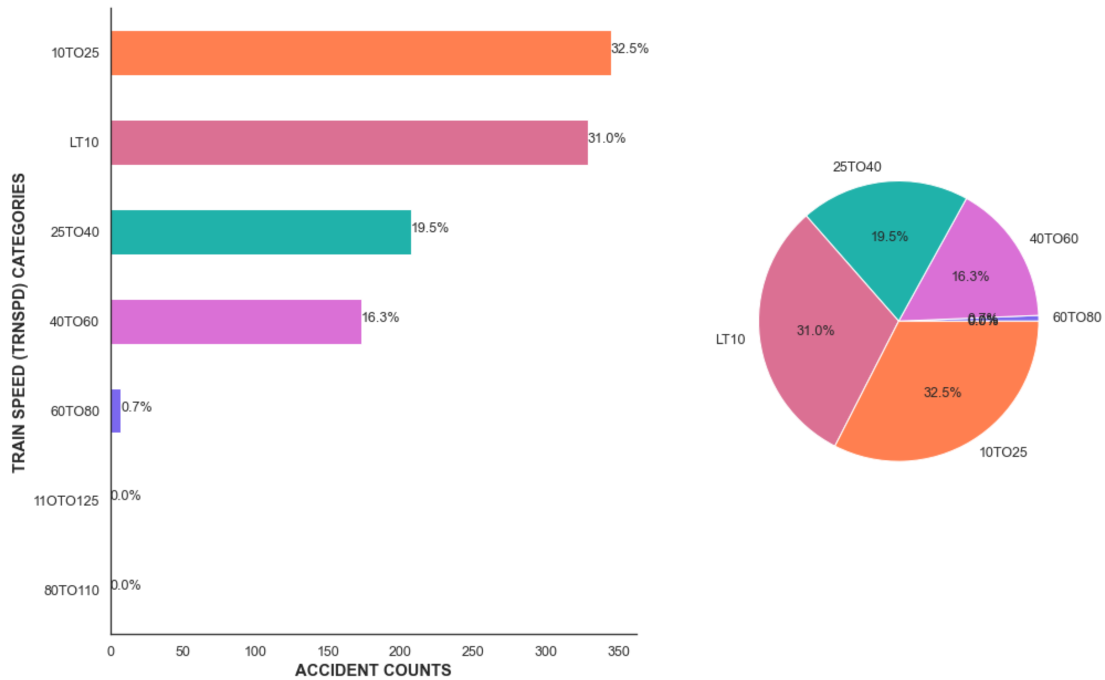


Figure 12. Distribution of Accidents Across Train Speed (TRNSPD) Categories

Table 12. Weather Categories

Weather Code	Weather Category
1	Clear
2	Cloudy
3	Rain
4	Fog
5	Sleet
6	Snow

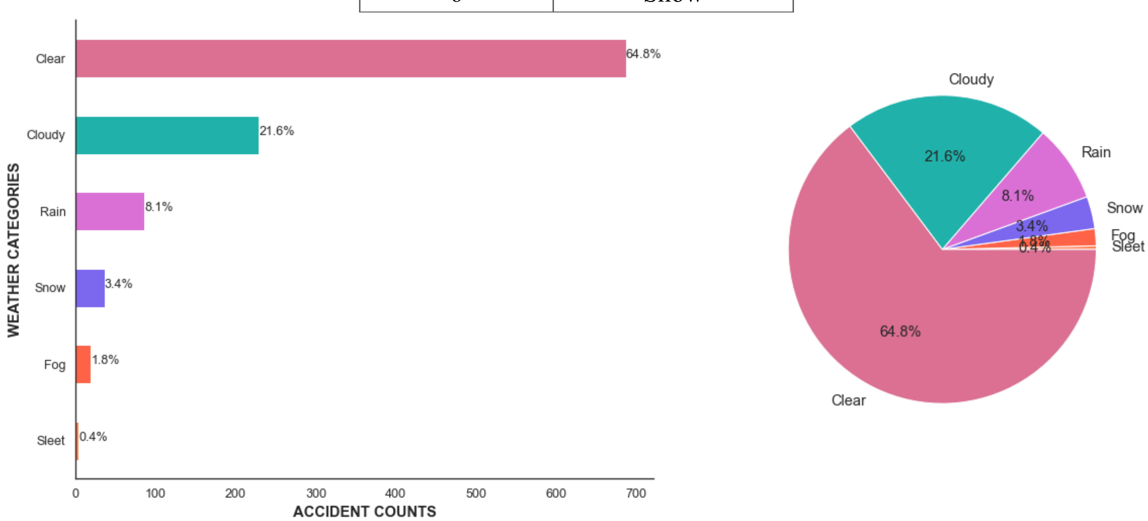


Figure 13. Distribution of Accident Categories Across Weather (WEATHER) Categories

Table 13. Temperature (TEMP) Categories

Temperature Category	Temperature Range
Very Cold	Min to 0F
Cold	0F to 40F
Moderate	40F to 60F
Warm	60F to 80F
Hot	80F to 100F
Very Hot	100F to Max

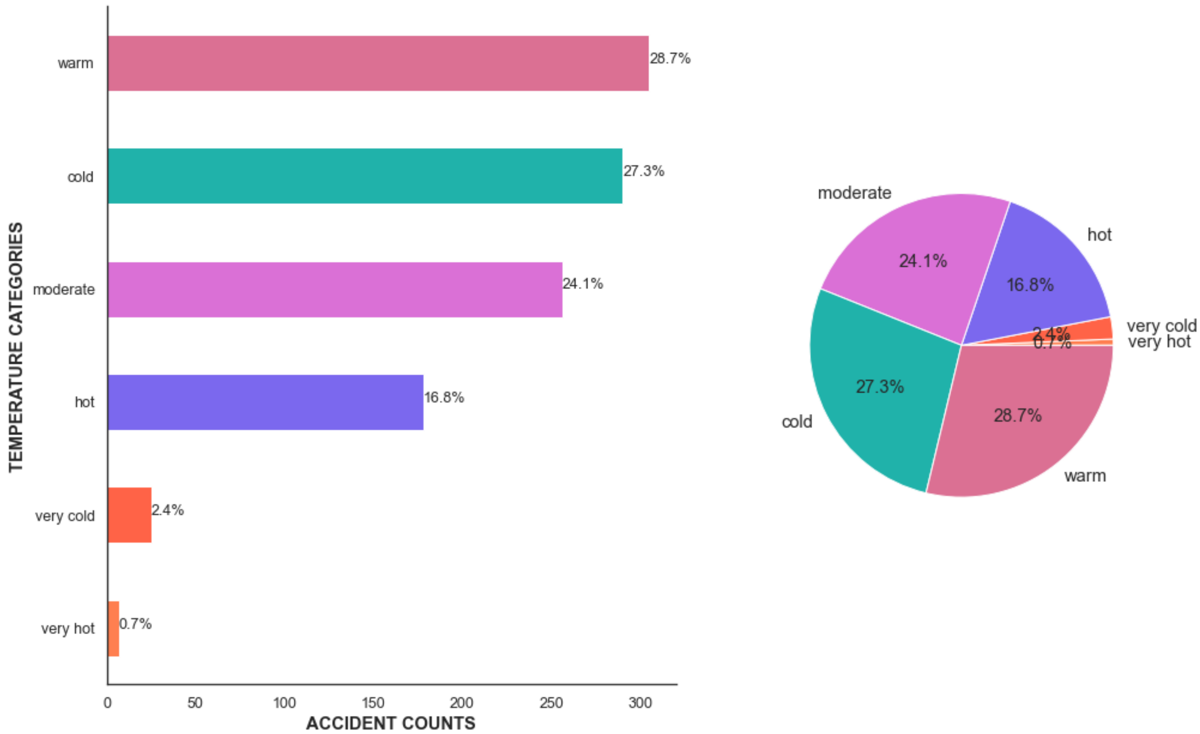


Figure 14. Distribution of Accidents Across Temperature (TEMP) Categories

Table 14. Visibility Categories

Visibility Code	Visibility Category
1	Dawn
2	Day
3	Dusk
4	Dark

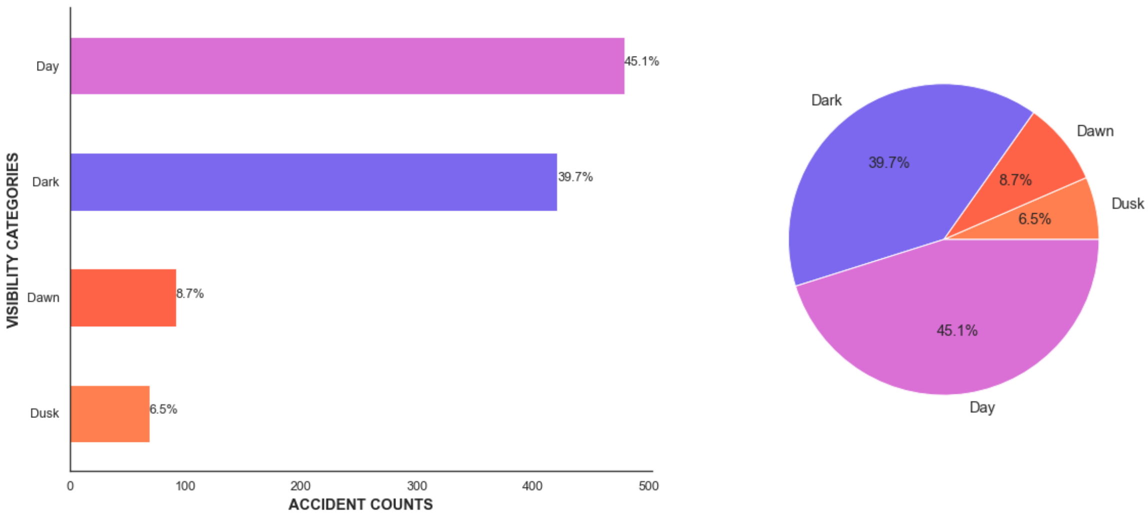


Figure 15. Distribution of Accidents Across Visibility (VISIBLTY) Categories

Table 15. Number of Loaded Freight Cars (LOADF1) Categories

Categories	# Loaded Freight Cars
LT35	Min to 35 cars
35TO70	35 to 70 cars
70TO100	70 to 100 cars
GT100	100 to Max cars

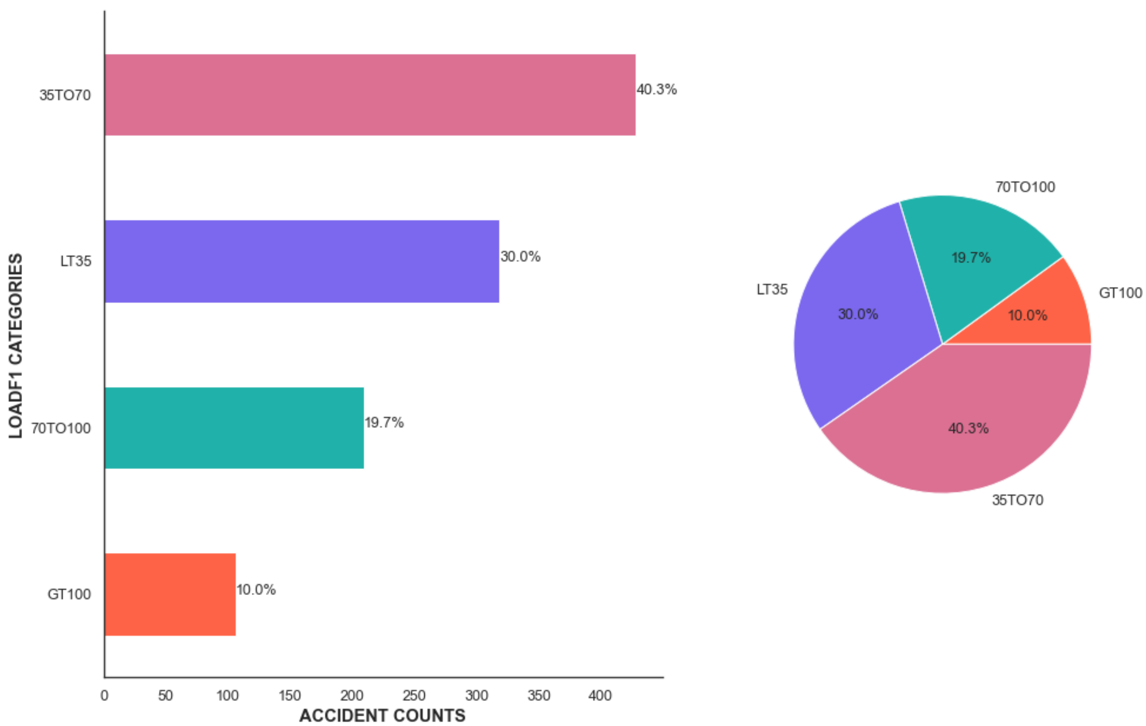


Figure 16. Distribution of Accidents Across Number of Loaded Freight Cars (LOADF1) Categories

Table 16. Total Number Of Cars (TOTALF1) Categories

Categories	# Loaded Freight Cars
LT50	Min to 50 total cars
50TO100	50 to 100 total cars
100TO150	100 to 150 total cars
GT150	Greater than 150 total cars

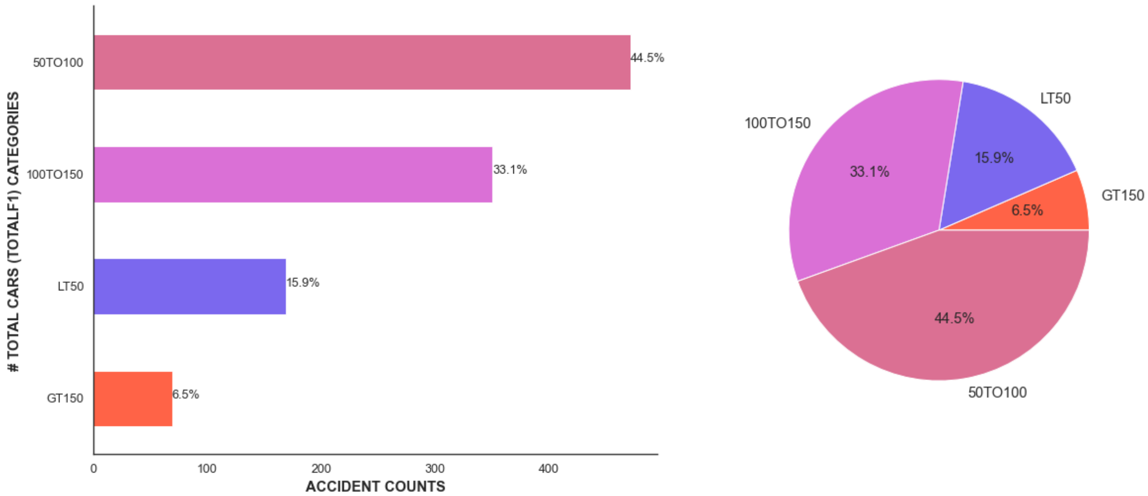


Figure 17. Distribution of Accidents Across Total Number of Cars (TOTALF1) Categories

Table 17. Gross Tons (TONS) Categories

Tonnage Category	Gross Tonnage (excluding power units)
LT5K	Min to 5000 tons
5TO10K	5000 to 10,000 tons
10TO15K	10,000 to 15,000 tons
GT15K	15,000 to Max tons

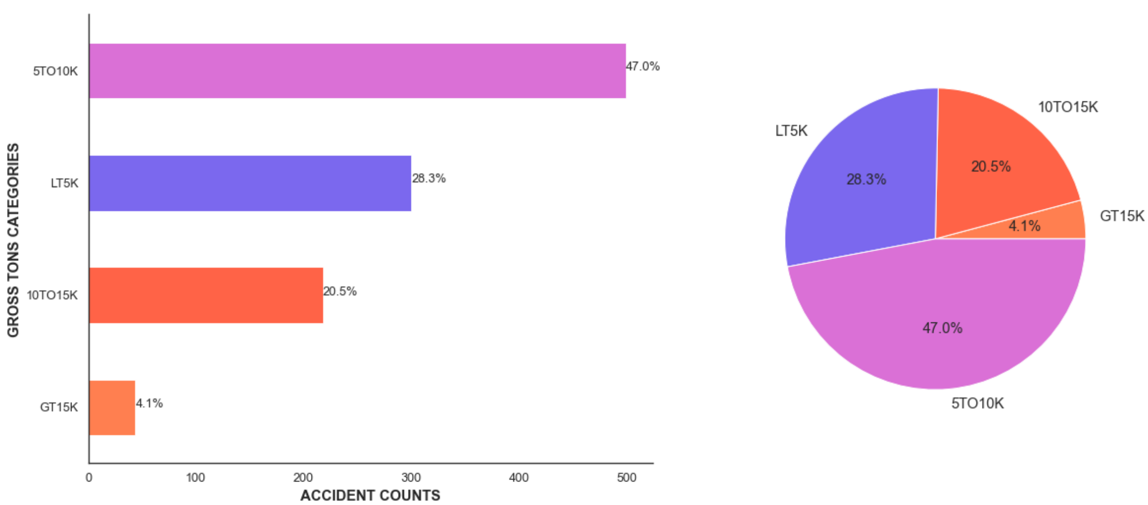


Figure 18. Distribution of Accidents Across Gross Tons (TONS) Categories

Table 18. Track Density (TRKDNSTY) Categories

Track Density Category	Track density in MGT
LT40	Min to 40 MGT track density
40TO80	40 to 80 MGT track density
80TO120	80 to 120 MGT track density
GT120	Greater than 120 MGT track density

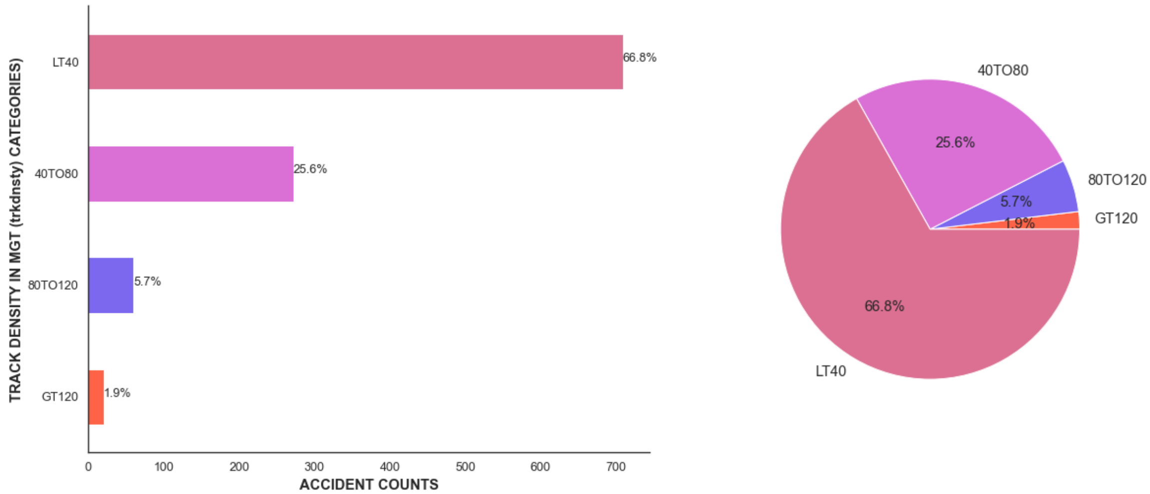


Figure 19. Distribution of Accidents Across Track Density in MGT (TRKDNSTY) Categories

3.5 Bayesian Network-based Railroad Accident Risk Model Implementation

The important features, assumptions, and limitations of the BN-based accident model are listed below.

1. The BN-based accident model is a tool that provides insight into the degree to which the causal factors influence railroad accidents, specifically derailments. This information will enable railroad companies to implement measures to monitor these factors to improve overall railroad safety and mitigate the risk of derailments.
2. RAIRS accident data between 2009 to 2018 is used to fit the BN-based accident model. Further, the Sort-C1 accident data considers freight trains operating on main tracks and carrying hazmat, with derailment type of accident attributed to track, roadbed, and structure-related causes for groups 03T, 04T, 05T, and 08T (Bing, et al., 2015).

The implementation of a BN consists of two components, structure learning and parameter learning.

Structure Learning – Structure learning involves the process of establishing the structure of the directed acyclic graph, which comprises nodes representing random variables and edges that establish the probabilistic causal relationship between the nodes. In this project, the parameters (i.e., risk factors) for constructing the BN were identified using the Random Forests method discussed in Section 3.3.2. Additionally, subject matter experts' recommendations were incorporated when finalizing the structures of the BNs implemented in this project.

Parameter Learning – This part of the BN implementation involves fitting the data to the Bayes model to learn the distributions for each of the nodes (i.e., parameters) in the network. Pomegranate uses the maximum likelihood estimate (MLE) method to develop either univariate or multivariate distributions for each of the nodes in the BN structure.

Bayesian Inferencing – Bayesian inferencing is the process of unearthing information once the data has been fit to an established network. Bayesian inferencing consists of calculating joint probabilities, performing predictive analytics, and using sensitivity analysis.

Joint Probability Estimation – This involves evaluating the total probability for each variable in a network or a subset of a network. The joint distribution for a BN is equal to the product of probabilities of the node given its parents:

$$P(X_1, X_2, \ldots .X_n = \prod_{i=1}^{n} P(X_i \mid Parents\ (X_i))$$

Predictive Analytics – An important feature of BNs is their ability to calculate the probability distribution over the unobserved variables (i.e., unknown parameters) given the evidence (i.e., known parameters). The more the observed variables, the higher is the confidence in the predicted values for the unobserved variables. Pomegranate uses the loopy belief algorithm to perform Bayesian inferencing.

Sensitivity Analysis – The likelihood of a railroad accident depends on all the contributing risk factors in the BN. However, different factors may have distinctive levels of influence on the railroad accident. Sensitivity analysis allows understanding the influence of the individual risk factors on the accident. These analyses help railroad companies and regulatory authorities develop safety policies that mitigate risk and reduce the number of accidents.

As mentioned earlier, the BN for probabilistic risk analysis of train derailments was implemented using pomegranate (Schreiber, 2018). The networks capture the causal as well as the statistical correlations between the risk factors and the primary accident causes. The following four networks were initially established by integrating the causal relationship learnt from the feature selection exercise in Section 3.3.2 and subject matter experts' knowledge.

3.5.1 Weather and Climate Factors

The accident data was fit to the BN for weather and climate related risk factors shown in Figure 20. The calculated marginal probabilities for the four track-related accident cause groups are shown in Figure 21. For the given data, none of the accident cause groups have a probability of more than 50 percent.

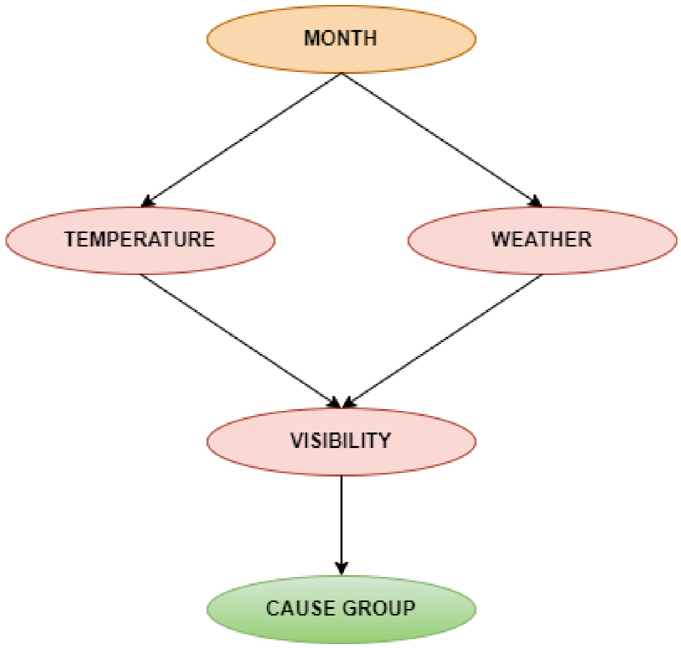


Figure 20. BN for Weather and Climate Risk Factors

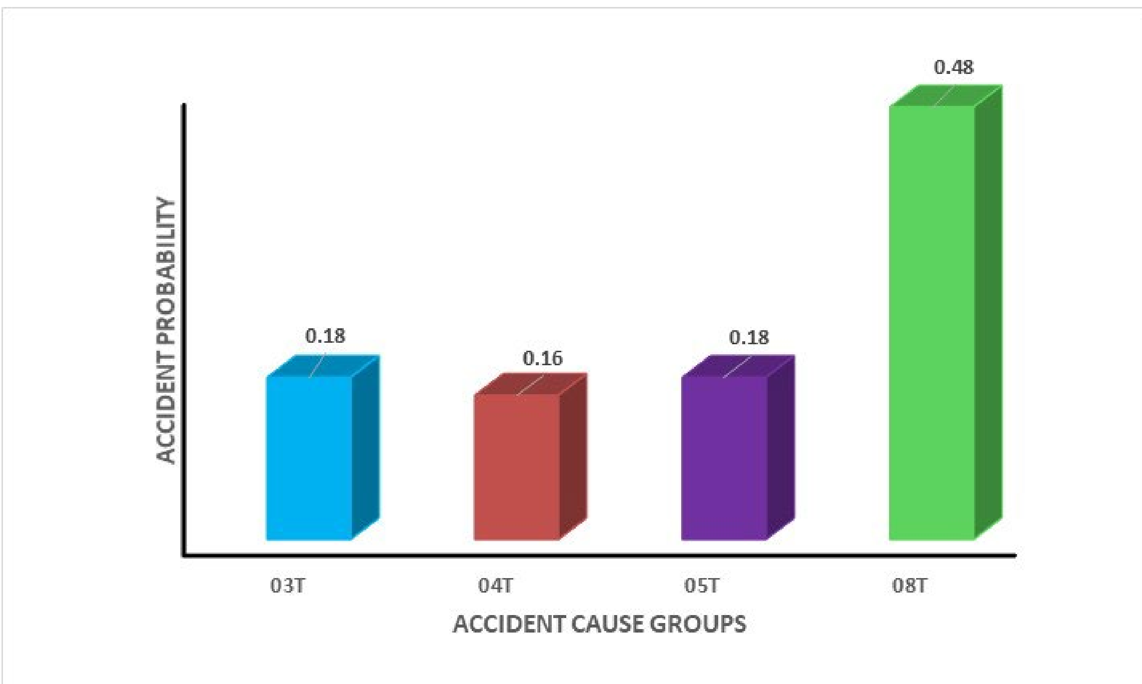


Figure 21. Marginal Probabilities for Weather and Climate Factors BN Fitted to the Sort C Accident Data (Derailment and Track Related Defects Only)

The predictive capabilities of the BN are illustrated through three observations listed in Table 19. In the first observation, the weather and the visibility are known, while temperature and the cause group are the unknown factors. With only two out of the four factors known, the BN cannot attribute the accident to one specific cause group. In the second observation, when information on temperature (TEMPERATURE=COLD) is added, the BN predicts the probable cause of an accident to broken rails and welds (08T). Likewise, when the temperature factor is set to very-hot (TEMPERATURE=VERY HOT) in the third observation, the BN network attributes the cause of the accident to buckled track (05T) with a 100 percent certainty. An important aspect to be noted here is that the BN has learned the underlying physics between cold temperatures causing material embrittlement and fracture (i.e., broken rails and welds) and hot temperatures causing track buckling through statistical correlations existing in the data.

Table 19. Accident Cause Group Predictions for Specific Weather and Visibility Conditions

Observations	Temperature	Weather	Visibility	Cause Group			
				03T	04T	05T	08T
#1	**UNKNOWN**	Clear	Day	18%	16%	18%	48%
#2	**COLD**	Clear	Day	21%	0%	0%	**79%**
#3	**VERY HOT**	Clear	Day	0%	0%	**100%**	0%

A sensitivity analysis for temperature was performed to understand the effect of varying temperature on accident causes. From the temperature sensitivity analysis plot in Figure 22, it is evident that cold temperatures lead to broken rails and welds (08T) and hot weather conditions cause track buckling (05T).

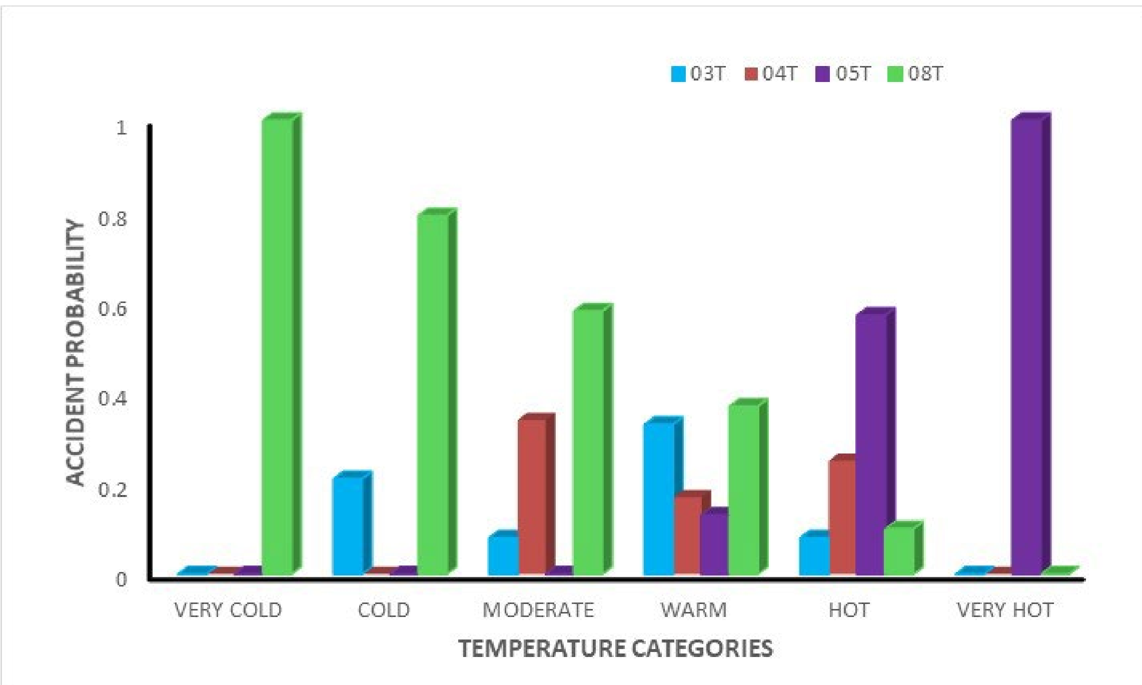


Figure 22. Sensitivity Analysis for Temperature - Plot of Accident Cause Group Probabilities Across Temperature Categories

3.5.2 Track Factors - Track Class and Track Annual Density

The BN for track-related risk factors is shown in Figure 23. The marginal probabilities for the four accident cause groups following fitting of the network to the accident data is shown in Figure 24. For the given data, none of the accident cause groups have a probability of more than 50 percent.

An example of the predictive capability of the network is presented in Table 20. In the first observation, the railroad company (Company-A) and the track density (TRK_DNSTY=LT40) are the observed variables, while the track class is the unobserved variable (unknown). With two out of the four variables known, the BN cannot attribute the accident to one specific cause group. In the second observation, in addition to the railroad company and the track density, the track class is also known (TRACK_CLASS = CLASS 1). After knowing the value for the track class, the BN increases the accident probability due to track geometry (04T) from 40 to 72 percent. The knowledge of increased accident risk on CLASS 1 tracks due to geometry issues (04T) can help railroad companies prioritize maintenance activities accordingly.

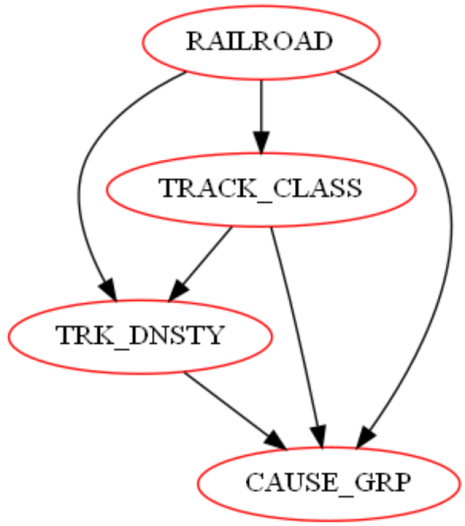


Figure 23. BN for Track Class and Track Annual Density

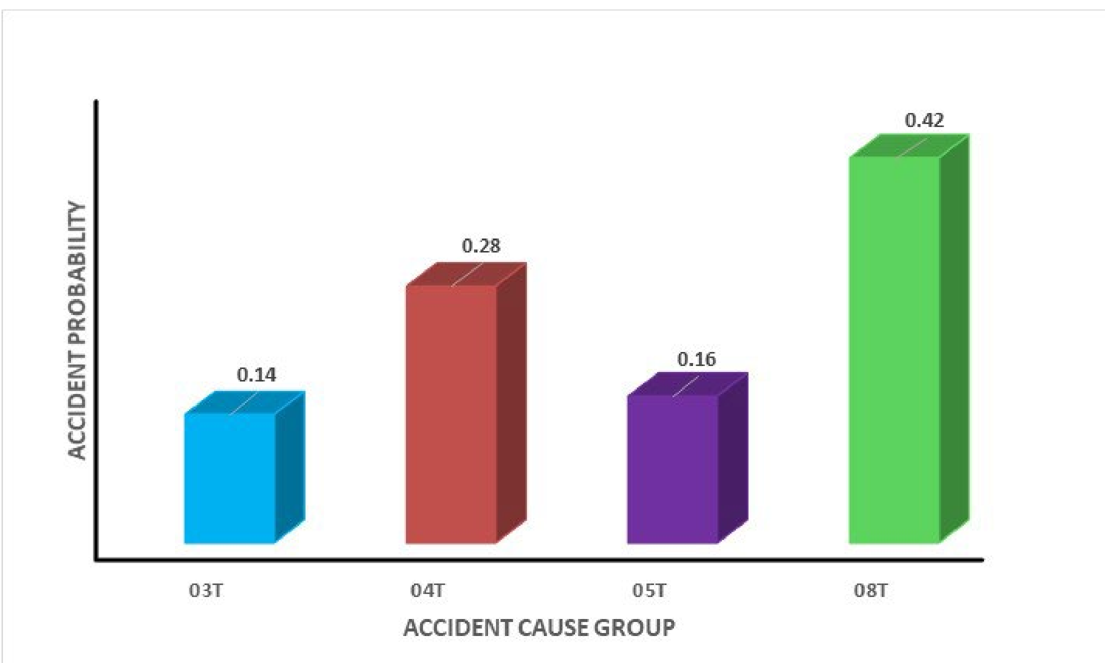


Figure 24. Marginal Probabilities for Track Class and Track Density BN Fitted to the Sort C Accident Data (Derailment and Track Related Defects Only)

Table 20. Accident Cause Group Predictions for Railroad Company-A

Observations	Track Class	Track Density	Cause Group			
			03T	04T	05T	08T
#1	**UNOBSERVED**	LT40	7%	**40%**	28%	25%
#2	CLASS 1	LT40	14%	**72%**	14%	0%

The plot of the sensitivity analysis for track class for railroad Company-A, operating on tracks with track density less than 40 MGT (TRK_DNSTY = LT40), is shown in Figure 25. This figure shows that accidents due to track geometry issues (04T) are more likely to occur on CLASS-1 and CLASS-2 tracks, whereas accidents due to broken rails and welds (08T) are more likely to occur on CLASS-3 and CLASS-4 tracks.

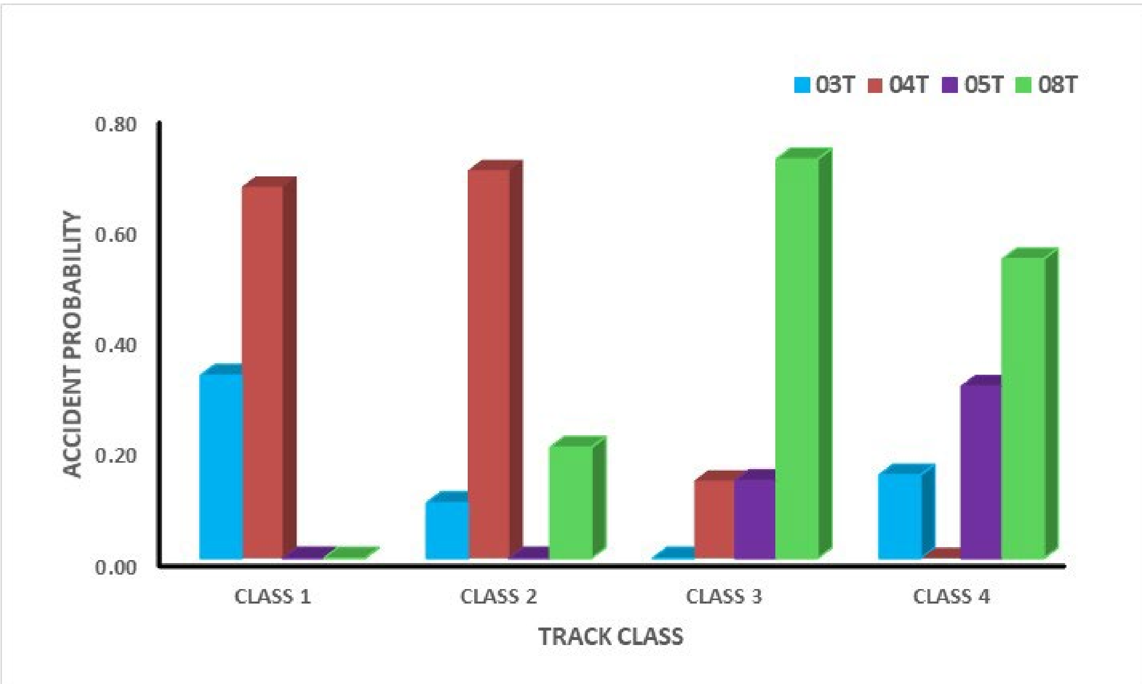


Figure 25. Sensitivity Analysis for Track Class - Plot of Accident Cause Group Probabilities Across Track Classes

3.5.3 Train Factors - Gross Tonnage and Freight Cars

The BN for gross tonnage and number of freight cars is shown in Figure 26. The marginal probabilities for the four accident cause groups following fitting of the network to the accident data is shown in Figure 27. For the given data, none of the accident cause groups have a probability of more than 50 percent.

Two examples illustrating the predictive capability of the network are presented in Table 21. In Observation #1 the railroad company (Company-A) and the number of loaded freight cars (LOADF1=LT35) are the observed (i.e., known) variables, and the total number of freight cars (TOTALF1) and the cause group are the unobserved (i.e., unknown) variables. The BN cannot attribute the cause of the accident to any specific cause group as only information about two out of the four variables is available. In the second observation the total number of cars is also known (TOTALF1=LT50) and the cause group is the only unknown variable. With three out of the four variables known, the BN attributes the cause of the accident to broken rails and welds (08T) with an 80 percent certainty. Likewise, in the second example (Observations #3 and #4), when information about three out of the four variables is available, the BN attributes the cause of the accident to track geometry (04T) with an 84 percent chance.

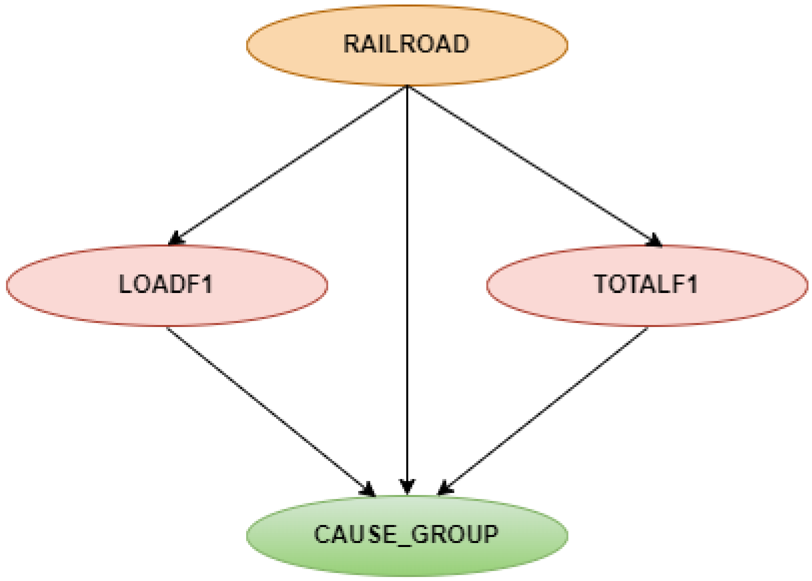


Figure 26. BN for Gross Tonnage and Number of Freight Cars

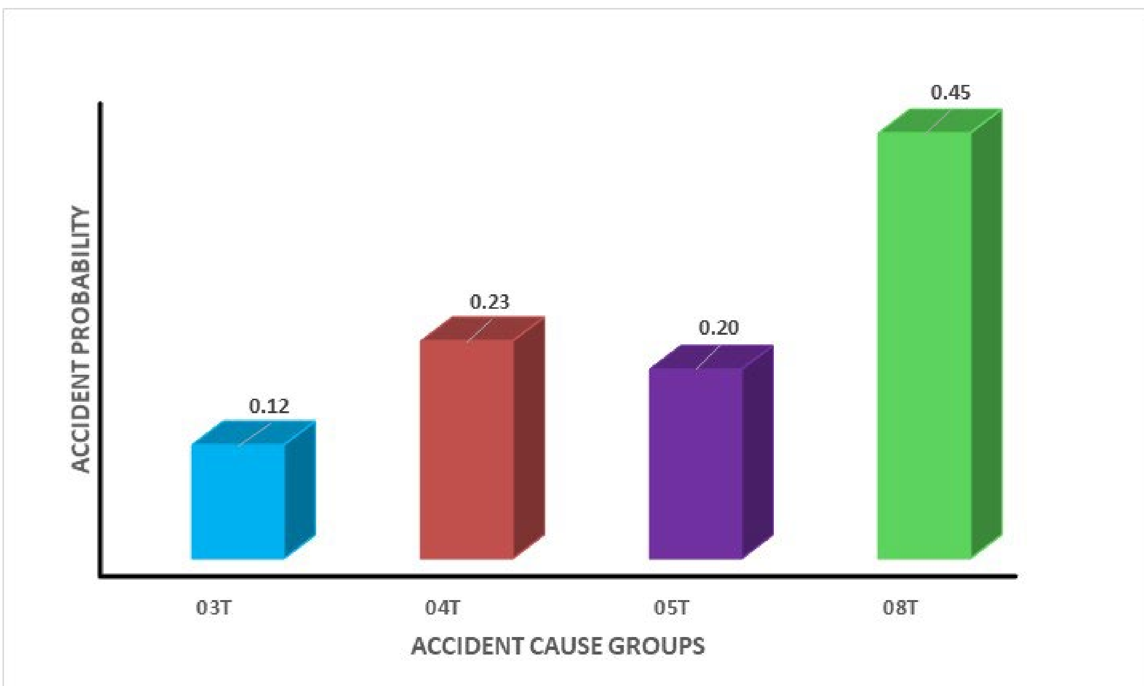


Figure 27. Marginal Probabilities for Tonnage and Freight Cars BN Fitted to the Sort C Accident Data (Derailment and Track Related Defects Only)

Table 21: Accident Cause Group Predictions for Railroad Company-A

Observations	Number of Loaded Cars LOADF1	Total Number of Cars TOTALF1	Cause Group			
			03T	04T	05T	08T
#1	LT35	**UNOBSERVED**	10%	33%	36%	**21%**
#2	LT35	LT50	0%	20%	0%	**80%**
#3	35TO70	**UNOBSERVED**	11%	**49%**	17%	23%
#4	35TO70	100 TO150	0%	**84%**	16%	0%

The plot of the sensitivity analysis for four railroad companies, with the total number of loaded cars less than thirty-five (LOADF1 = LT35) and the total number of cars less than fifty (TOTALF1=LT50) is shown in Figure 28. For the given number of loaded cars and total number of freight cars, the cause of an accident is attributed to broken rails and welds (08T) for three out of the four railroads. For railroad Company-B the BN attributes the cause to track geometry issues (04T) and broken welds and rails (08T) with equal probability of 50 percent.

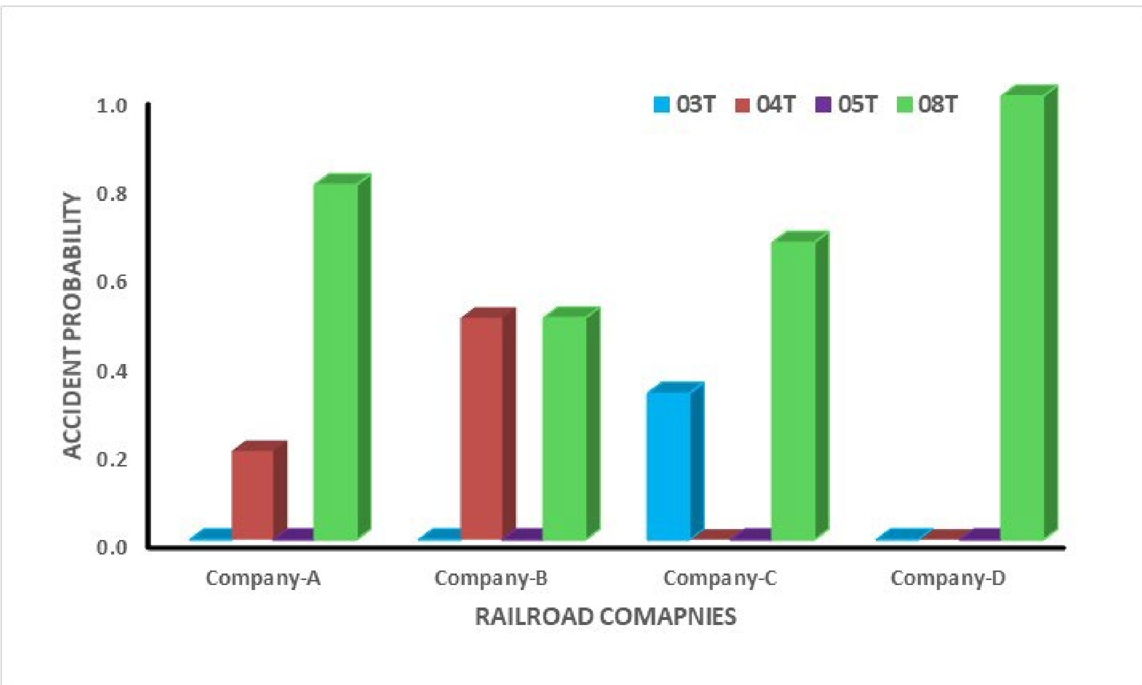


Figure 28. Sensitivity Analysis for Railroad Companies - Plot of Accident Cause Group Probabilities Across Four Railroad Companies

3.5.4 Unified Bayesian Network – Track and Train Factors

Next the rail operations, train composition, and tonnage networks are grouped into a unified network. Whereas the weather and climate parameters provide a good understanding of the causal relationship, these factors are beyond the control of railroad companies and therefore were not incorporated into the unified network. The unified BN for derailment risk analysis is presented in Figure 29.

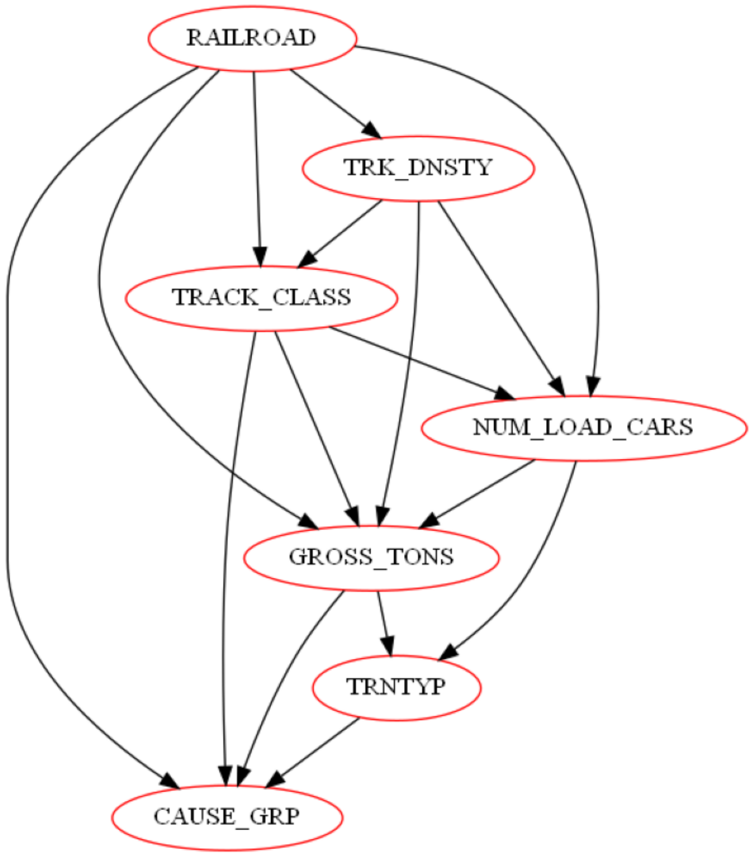


Figure 29. Unified BN for Rail Operations, Tonnage, and Composition

In the unified BN a term has been introduced for train type, which identifies the train as either unit train or manifest train. Unit trains were identified using the ruleset described in Section 2.1.3. Review of the model predictions found that train type did not have significant influence on the predicted accident cause. This is suspected to be due to the limited scope of the current model, which is built upon a subset of Sort C data comprising train derailment accidents caused by four specific cause groups: Wide Gauge (03T), Track Geometry (04T), Buckled Track (05T), and Broken Rail or Welds (08T). This limits the number of unit train accidents in this data set and the ability to identify differences in accident properties. To address this limitation, the team recommends further development of the model to include additional accident cause code groups, other accident types outside of train derailment, and a larger time period.

3.6 Bayesian Network Validation

This section discusses the details of the validation of the unified BN discussed in Section 3.5. Researchers used RAIRS accident data for the years 2019 and 2020 to validate the BN. A total of 2750 accidents were recorded in 2019 and another 2214 accidents were recorded in 2020. The following three filters were applied to align the data with the 2008-2019 data set used for training the BN.

1. Sort C filter (freight train on main track consisting of hazmat cars)
2. Derailment type of accident

3. Accident cause code groups (03T, 04T, 05T, and 08T)

After application of the above filters, the data was reduced to 4 accidents from 2019 and 20 accidents from the year 2020. The data from these 24 accidents was input to the unified BN to predict the accident cause group for each of the accidents. Table 22 lists a summary of the accuracy of the predictions from the unified BN. The derailment risk model achieves a good accuracy for the 03T, 04T, and 08T groups. The lower accuracy for the 05T group is attributed to the small number of accidents belonging to that group in Sort C (i.e., the sort used to train the BN). However, for a training dataset comprising 296 accidents, an overall accuracy of 79 percent is reasonably good. The accuracy of the model can be improved by augmenting the Sort C accident data.

Table 22. Verification Results

Cause Group	Number of Accidents	Accidents Accurately Predicted	Accuracy	Sort-C1 Accidents by Cause Group
Wide Gauge (03T)	1	1	100%	39
Track Geometry (04T)	11	8	73%	66
Buckled Track (05T)	3	1	33%	49
Broken Rail or Welds (08T)	9	9	100%	142
Overall	**24**	**19**	**79%**	**296**

4. Conclusions and Recommendations

4.1 Conclusions

The objective of this research project was to develop a predictive risk model for the release of hazardous material transported in unit trains using data science techniques for available rail accident and traffic data. This work builds upon previous research (Bing, et al., 2015) which examined the causal sequence of events that can lead to a rail accident and used historical accident record and rail traffic data to define conditional probabilities of occurrence and thereby predict the risk of a hazmat release.

The research team reviewed sources of accident data and traffic data to inform the development of a risk model for hazmat transportation by rail. The team concluded that due to the limitations of the available traffic data, the research would focus on building a Bayesian-based accident model using RAIRS accident data.

The most detailed traffic data, the STB CCWS, provides information about car and commodity movements, but does not provide information about the train in which the car was transported, limiting its ability to address the central question of the research. Other traffic data reviewed, including Class I railroad R-1 reports (Surface Transportation Board, 2021a) and AAR's Ten-Year Trends (Association of American Railroads, 2020), provide annual statistics on freight transport, including unit train miles, but do not distinguish hazmat unit trains from non-hazmat unit trains.

Consequently, there was no means of calculating risk, even in the simple terms of accident likelihood per unit distance travelled by hazmat unit train versus hazmat manifest train. Therefore, the research focused on building a predictive model based solely on the RAIRS accident data. The aim of the model was to predict the cause of an accident, given information about the accident, based on trends and associations derived from 10 years of historical accident records. The trends and associations derived from the accident data could then indicate differences in unit train accidents compared to manifest trains and speak to the central question regarding relative risk of hazmat transportation.

Three different BNs were implemented to study the causal relationships between weather, track, and train related risk factors and the primary causes leading to railroad accidents. The primary causes were selected based on the risks/consequences they posed. These causes were then categorized according to the grouping suggested in the work by Bing et. al (2015). A Random Forest algorithm was used to select the risk factors. The risk factors in continuous data form were then converted into discrete categories by studying their distributions. The capabilities of the BN-based accident risk model, such as predictive analytics, sensitivity analysis, and estimation of marginal and joint probability distributions, were demonstrated through the implementation of three separate BNs for weather, track, and train related factors. The final network integrated the train and track networks into one single network, and included a variable for train type (i.e., unit train). Hazmat unit trains were identified in the RAIRS database based on the number of hazmat cars (≥70) and the number of buffer cars (≤5).

Review of the model predictions found that train type did not have significant influence on the predicted accident cause. This is suspected to be due to the limited scope of the current model, which is built upon a subset of data comprising hazmat freight trains travelling on main track, where a derailment occurred, and the cause was one of four specific cause groups: Wide Gauge (03T), Track Geometry (04T), Buckled Track (05T), and Broken Rail or Welds (08T). This limits the number of unit train accidents in this data set and the ability to identify differences in accident properties.

Accidents from 2019 and 2020 were used to validate the integrated BN and data filters used for developing the Sort-C1 were applied on this accident database. In total, 24 accidents were selected. An important aspect of the validation is that the BN has not seen this data previously. The risk factors from the 24 accidents were provided as inputs to the BN and the network was tasked to make blind predictions on the cause for these accidents. The integrated network predicted with 100 percent accuracy for accident cause groups 03T and 08T and 79 percent accuracy for 04T. The lower accuracy for cause group 05T is attributed to the small amount of training data in the Sort-C1. However, the overall accuracy of the network predictions was about 79 percent, which is promising. Increasing the training data pool (Sort C1) is expected to improve the accuracy of the network.

4.2 Recommendations

To aid future research on the risk of hazmat transportation by unit train, improvements in accident and traffic data collection are needed.

In terms of accident data, the research team recommends that a new field be added to the FRA RAIRS database, indicating whether the train is a unit train. During the research, significant variation was observed in how a unit train is defined. Therefore, such a change would require adoption of a standard definition of a unit train.

It may not be feasible to collect the level of information required to fully incorporate traffic data into the architecture of the Bayesian-based accident model, even in a geographically limited model. This would require that all successful train journeys over a given time-period are added into the database alongside those which ended in an accident. However, for more traditional statistical methods, improvements could be made to identifying hazmat unit train traffic. Schedule 755 of the R-1 reports (Railroad Operating Statistics) submitted annually to the STB (Surface Transportation Board, 2021a) includes data on total train-miles and car-miles by train type, which includes unit trains. Therefore, the team recommends that the collection of railroad operating statistics be expanded to provide total train-miles and car-miles by train type transporting hazmat.

The BN-based accident model has demonstrated potential to accurately predict accident cause, given information about the train and track. The team recommends that consideration is given to further developing the capabilities of the BN-based accident risk model by incorporating additional causes groups (e.g., mechanical and electrical, and human causes). In addition, inclusion of track maintenance inspection and repair data could enhance causal relationship learning.

5. References

Association of American Railroads (2020). Railroad Ten-Year Trends 2009-2018. Washington D.C.: AAR Policy and Economics Department.

Bing, A., Dizon, A., Brickett, J., Papson, A., & O'Rourke, L. (2015). Risk Evaluation Framework and Selected Metrics for Tank Cars Carrying Hazardous Materials (Report No. DOT/FRA/ORD-15/07). Federal Railroad Administration.

BNSF (2022). Glossary of Railroad Terminology & Jargon. [Accessed 13 May 2022]

Cheng, C.-W., Leu, S.-S., Lin, C.-C. & Fan, C. (2010). Characteristic analysis of occupational accidents at small construction enterprises. *Safety Science, 48*(6), 698-707.

Canadian National (2001). Annual Report. Montreal: Canadian National Public Affairs.

Canadian Pacific (2019). Unit Train Services Tariff 5. Calgary: Specialist - Tariff Policy Canadian Pacific.

CSX (2022). Railroad Dictionary. [Accessed 13 May 2022]

Fan, C.-F. & Yu, Y.-C. (2004). BBN-based software project risk management. *Journal of Systems and Software, 73*(2), 193-203.

Hänninen, M. (2014). Bayesian networks for maritime traffic accident prevention: Benefits and challenges. *Accident Analysis & Prevention, 73*(December), 305-312.

Heckerman, D., Mamdani, A., & Wellman, M. P. (1995). Real-World Applications of Bayesian Networks. *Communications of the ACM, 38*(3), 24-26.

Kansas City Southern (2022). Terms & Definitions. [Accessed 13 May 2022]

Liu, X. (2017). Risk Comparison of Transporting Hazardous Materials in Unit Trains Versus Mixed Trains. *Transportation Research Record: Journal of the Transportation Research Board, 2608*(1), 134-142.

Li, W., Roscoe, G., Zhang, Z., Saat, R., & Barkan, C. (2018). Quantitative Analysis of the Derailment Characteristics of Loaded and Empty Unit Trains. *Transportation Research Record: Journal of the Transportation Research Board, 2672*(10), 156-165.

Camino López, M. A., Ritzel, D. O., Fontaneda, I., & Gonzalez Alcantara, O. J. (2008). Construction industry accidents in Spain. *Journal of Safety Research, 39*(5), 497-507.

Luu, V. T., Kim, S.-Y., Tuan, N. V., & Ogunlana, S. O. (2009). Quantifying schedule risk in construction projects using Bayesian belief networks. *International Journal of Project Management, 27*(1), 39-50.

Martín, J. E., Rivas, T., Matías, J. M., Taboada, J., Argüelles, A. (2009). A Bayesian network analysis of workplace accidents caused by falls from a height. *Safety Science, 47*(2), 206-214.

National Transportation Safety Board (2020). Placement of DOT-111 Tank Cars in High Hazard Flammable Trains and the Use of Buffer Cars for the Protection of Train Crews (Investigation No. RRD19FR007/ RRD20FR002). Washington, DC: NTSB.

Norfolk Southern (2019). Freight Tariff NS 6004-D. Pricing Services, Norfolk Southern Railway Company and Consolidated Subsidiaries.

Salvatier, J., Wieki, T., & Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science, 2:e55*.

Schmidt, B. (2015). Breaking the code. Trains Magazine.

Schreiber, J. (2018). Pomegranate: fast and flexible probabilistic modeling in python. *Journal of Machine Learning Research, 18*, 1-6.

Starr, J. T. (1976). Evolution of the Unit Train: 1960-1969. Chicago: University of Chicago Press.

Surface Transportation Board (2021a). Annual Report Financial Data. [Accessed 18 May 2022]

Surface Transportation Board (2021b). Carload Waybill Sample. [Accessed 13 May 2022]

Union Pacific (2022). Rail Within Reach. [Accessed 13 May 2022]

Uusitalo, L., 2007. Advantages and challenges of Bayesian networks in environmental modelling. *Ecological Modelling, 203*(3-4), 312-318.

Zhang, F. L., Xiong, H. B., Shi, W. X. & Ou, X. (2016). Structural health monitoring of Shanghai Tower during different stages using a Bayesian approach. *Structural Control and Health Monitoring, 23*(11), 1366-1384.

Zhang, L., Wu, X., Skibniewski, M., Zhong, J., & Lu, Y. (2014). Bayesian-network-based safety risk analysis in construction projects. *Reliability Engineering & System Safety, 131*(November), 29-39.

Appendix A.
Class I Unit Train Definitions

A *Trains* magazine article provides insight on the meaning behind the train number assigned to Class I railroad operators and their subsidiaries (Schmidt, 2015). The train number is documented in the RAIRS database system (under the field title "TRNNBR"), which provides an alternate approach to identifying unit trains. However, as it is limited to Class I railroads and their subsidiaries, it was not adopted for the purposes of this research, which instead used a consistent rule for train length and number of buffer cars to identify unit trains. The train number is prefix dependent on the material being transported, and those which describe unit trains are summarized in Table 23.

Table 23. Train Number Prefix Definitions – Unit Trains

Class I Railroad Operator	Train Number Prefix	Description
BNSF	C	Loaded coal trains
	E	Empty coal trains
	G	Unit grain trains
	U	Unit trains
	X	Empty grain trains
CN	B3	Unit potash trains
	B7	Unit potash trains
	C7	Coal trains
	G8	Grain trains
	S7	Unit sulphur trains
	U7	Other unit trains
CP	300	Grain trains
	600	Unit trains
	800	Coal trains
KCS	C	Unit coal
	G	Unit grain
	L	Coal
	O	Unit aggregate
	U	Other unit
CSX	E	Empty unit trains
	G	Grain trains
	K	Various unit trains
	N	Loaded coal trains
	T	Loaded coal trains
	U	Loaded coal trains
	V	Grain trains
NS	400 or X	Coal and coke trains
	500 or Y	Coal and grain trains
	600 or Z	Coal and unit trains
	700 or Q	Coal trains
	800 or S	Coal trains
UP	C	Coal trains
	G	Grain trains
	O	Ore, crude oil trains
	U	Other unit trains

Table 24. Class I Railroads and Subsidiary Company – Acronyms RAIRS

Railroad	Unit Train Definition
BNSF	BN, LAJ, BNSF
CSX	CSX, BOCT, CARR, AWRY, DMRR, RFP, SIRT and TTIS
NS	NS, SOU, CNTP, AGS, CGA, CRSH
CPRS	CPRS, CP, SOO, DH, DME
CN	CN, GTW, IC, CEDR, CC, WC, EJE, DWP.
KCS	KCS, GWWR, GWWE, TM
UP	UP, ALS, CCT, CWI, SJ&GI, TCT, PTO

Further definitions were obtained from websites and reports published by Class I railroads, as shown in Table 25. The variations illustrate the lack of a consistent definition across the rail industry. In some cases, total number of cars is provided and ranges between 50 and 90.

Table 25. Unit Train Definitions as Provided by Railroads

Railroad	Unit Train Definition	Reference
BNSF	Non-stop service between a single origin and destination	(BNSF, 2022)
CSX	A train operating generally intact between point of origin and final destination, normally hauling a single bulk commodity, composed of like cars, equipped with high-tensile couplers	(CSX, 2022)
NS	A railway train with a minimum of 50 cars that is permitted and approved by NS to move in Merchandise Unit Train service and where all of the customer's railcars are moving from a single origin to a single destination	(NS, 2019)
CPRS	Unit Train or Solid Train means a physically consecutive and connected set of at least 80 cars tendered for movement together, unless otherwise stated in your contract	(CP, 2019)
CN	A train with a fixed, coupled consist of cars operated continuously in shuttle service under load from origin and delivered intact at destination and returning usually for reloading at the same origin	(CN, 2001)
KCS	Train that carries the same cargo (opposite of Manifest train); a train whose cars all carry the same commodity, such as grain or oil	(KCS, 2022)
UP	Unit trains transport more than 90 rail cars of one type of freight in one car type for one destination	(UP, 2022)

Abbreviations and Acronyms

ACRONYM	DEFINITION
AAR	Association of American Railroads
BN	Bayesian Network
CCWS	Confidential Carload Waybill Samples
DAG	Directed Acyclic Graph
FRA	Federal Railroad Administration
hazmat	hazardous material
HHFUT	High-Hazard Flammable Unit Train
ML	Machine Learning
MLE	Maximum Likelihood Estimate
MSE	Mean Squared Error
NTSB	National Transportation Safety Board
RAIRS	Railroad Accident/Incident Reporting System
SME	Subject Matter Expert
STB	Surface Transportation Board