U.S. Department of Energy

Fire Protection Data Trends, Calendar Years 2015-2019

December 2021

(Page intentionally left blank)

Table of Contents

(Page intentionally left blank)

Figures

Tables

Acronyms

ANL	Argonne National Laboratory
CAIRS	Computerized Accident and Incident Reporting System
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DOE	U.S. Department of Energy
ES&H	Environment, Safety and Health
FPDB	Fire Protection Database
FPAS	Fire Protection Annual Summary
NREL	National Renewable Energy Laboratory
ORPS	Occurrence Reporting and Processing System
OSHA	Occupational Safety and Health Administration
PCA	Principal Component Analysis
PPPL	Princeton Plasma Physics Laboratory
TF-IDF	Term Frequency Inverse Document Frequency
TSNE	T-distributed Stochastic Neighbor Embedding

(Page intentionally left blank)

Executive Summary

The U.S. Department of Energy (DOE) Office of Environment, Safety, and Health (ES&H) Reporting and Analysis within the Office of the Associate Under Secretary for Environment, Health, Safety and Security (AU) develops trending analyses of fire protection data to provide DOE Management, Program Offices, workers, fire protection personnel and other stakeholders an evaluation of DOE-wide performance and insights.

The U.S. Department of Energy Annual Fire Protection Data Trends for Calendar Years (CY) 2015-2019 presents the results of the observed trends of fire protection data at DOE operations during the 5-calendar year period of 2015-2019. This report analyzes information from over 59 DOE organizations that submit data to the DOE's Fire Protection Database (FPDB) as well as those that submit fire protection information to the Occurrence Reporting and Processing System (ORPS). The information is reported to DOE in accordance with DOE Order 231.1B, Environment, Safety, and Health Reporting. The information has been analyzed to provide a measure of DOE's fire protection safety performance.

Analysis of the collected fire protection data between CY2015-2019 indicates that the number of fire protection incident trends are flat and the normalized trends (i.e., per 200K work hours) are decreasing. However, the analysis found that the trends for total monetary losses are increasing. This may indicate that incidents are becoming less frequent but resulting in higher monetary losses. Vegetation or brush fire/smoke incidents and design/material related causes were found to have the highest increasing trends. The evaluation utilized natural language processing and machine learning text clustering methods to analyze the text data of the fire protection loss reports. The results identified that the most common fire protection incidents resulting in losses greater than $10K were related to the following:

- Vegetation related fires caused by lightning strikes, downed power lines, cigarettes, equipment failures (e.g., electrical arcing)
- Cold weather-related incidents resulting in building flooding damage due to frozen fire protection system pipes
- Equipment failures (e.g., transformers, fans, capacitors, modulators and heating, ventilation, and air conditioning)
- Chemical reactions within "fume hoods"
- Vehicle accidents
- Kitchen and analytical microwave smoke/fire incidents due to equipment failure or operator error

The analysis also evaluated FPDB fire department response data at four sites. However, due to limitations of the data no insights could be obtained. This limitation could be addressed by improving reporting of the fire department response data to ensure data consistency and reliability. AU is working to improve data collection, data quality, ability to measure trends and identify areas of improvements and potential leading indicators that could enhance learning and safety across DOE. Improvements will be coordinated with the users of the DOE's FPDB as part of a separate initiative.

(Page intentionally left blank)

1 Introduction

The U.S. Department of Energy (DOE) collects operating experience information in several databases to identify insights for operational improvements. This report discusses trends[1] related to fire protection data with the objective of identifying trends or issues that may be of interest to decision-makers and safety personnel.

This report utilizes fire protection related data from the Fire Protection Database (FPDB) and the Occurrence Reporting and Processing System (ORPS); and limited data from the Computerized Accident and Incident Reporting System (CAIRS) from calendar years (CY) 2015 to 2019 and is presented in CY. Where applicable, this report illustrates the differences between normalized (per 200,000 work hours) and non-normalized data over the five-year period, and also discusses limitations of the data such as potential reporting differences between the sites. This report analyses text of the fire protection reports using machine learning to identify frequent and recurrent terms, words, and topics.

1.1 Report Organization

This report is organized into 4 sections. Section 1 describes the background, content, and organization of this report. Section 2 discusses DOE's FPDB reports losses and trends. Section 3 discusses the Occurrence Reporting and Processing System (ORPS) fire protection related reports. Section 4 discusses FPDB fire department response trends. Section 5 discusses conclusions.

1.2 Fire Protection Annual Summary (FPAS) and Fire Protection Database (FPDB)

The DOE Order 231.1B, Environment, Safety, and Health Reporting, requires AU to develop the Fire Protection Annual Summary (FPAS) report. To accomplish this, the field organizations responsible for maintaining property enter fire related reports and reporting data into the DOE FPDB. The FPAS contains summary data for the preceding year as submitted to Headquarters, in accordance with the Annual Fire Protection Summary Information Reporting Guide.

Managed by the Office of Environment, Safety and Health (ES&H) Reporting and Analysis, the FPDB supports DOE efforts to gather and analyze fire protection related information such as reports and monetary losses, and fire department program costs and response data. The monetary loss reports include those related to fire protection losses which are divided into fire losses and non-fire losses and defined in the FPAS. The database is used to develop the FPAS report and analyzing data trends.

The FPAS defines fire protection loss as all damage or loss sustained as a consequence of fire events, including non-fire events. Fire loss is defined as all damage or loss sustained as a direct consequence of (and following the outbreak of) a fire. Non-Fire Loss is defined as all damage sustained as a consequence of non-fire events involving fire protection systems, such as leaks, spills, and inadvertent releases.

1.3 Occurrence Reporting and Processing System (ORPS)

The requirements of ORPS requirements are contained in DOE Order 232.2A Chg1 (MinChg), Occurrence Reporting and Processing of Operations Information. Within 2 hours of discovery of an event or condition, DOE site personnel must determine if the event or condition meets the threshold for reporting into the ORPS database, and if so, which reporting criteria applies. When the reporting criteria

[1] Calculated using a linear trendline.

is met, sites submit occurrence information such as title, description, dates, organization, reporting criteria, condition, cause codes, actions taken, and corrective actions.

DOE AU staff review all ORPS reports and categorizes each report by assigning Headquarters keywords (see Table 1). For keywords within the EH Categories category only the report's most applicable keyword is assigned. For keywords within the Quality Assurance category at least one keyword is assigned to a report, but multiple keywords may be assigned if applicable. For other categories an ORPS report can have as many keywords as are applicable. For example, an ORPS report that is related to category 3, Fire Protection and Explosive Safety, can have multiple keywords assigned such as Facility Fire and Fire Suppression Actuation, for a report that is related to a fire where a fire suppression system actuated. The ORPS data can then be used to perform trending analyses, evaluate occurrence causes, and identify lessons learned to improve operational safety.

DOE updated Order 232.2A on January 17, 2017, which is within the trending period of this report. Major changes in the updated version of the order include revision to reporting levels and requirements, making reporting of informational events optional at the discretion of the Program Office. The 2017 revision also retired several keywords, which are noted in Table 1, and created the following new keywords:8M-Chemical Safety, 8N-Laser Safety, 8O-Consturction Demolition Safety, 8P-Hoisting/Rigging Incident, 8Q-Forklift/Hand Truck Incident, 8R-Excavations/Penetrations, 8S-Landscaping/Mowing and 8T-Beryllium Incident.

Table 1: ORPS Event Oriented Headquarter Keywords

1. Work Planning and Control Deficiencies		2. Environmental	3. Fire Protection and Explosives Safety	4. Instrumentation and Controls	5. Mechanical/Structural
A. Inadequate Conduct of Operations (Retired) B. Loss of Configuration Management/Control C. Violation of Authorization Basis Elements D. Missed/Late Surveillance E. Facility Operations Procedure Noncompliance F. Training Deficiency G. Inadequate Procedure H. Inadequate Safety Analysis/USQs/TSRs I. Safety System Actuation/Evacuation J. Criticality Procedure Noncompliance K. Lockout/Tagout Noncompliance (Electrical)	L. Lockout/Tagout Noncompliance (Other) M. Inadequate Job Planning (Electrical) N. Inadequate Job Planning (Other) O. Inadequate Maintenance P. Inadequate Oral Communication Q. Personnel Error R. Management Issues S. Incorrect/Inadequate Installation T. Willful Violation U. Unplanned Interruption of Operations	A. Radioactive Release B. Underground Storage Tank Release C. Compliance Notification (from regulator with a violation) D. Compliance Notification (to regulator without a violation) E. Hazardous Material Release F. Potable Water Release	A. Fire Protection Equip Degradation B. Fire Suppression Actuation C. Facility Fire D. Explosives Safety Issue E. National Fire Protection Association (NFPA)/Code/Fire Protection Issues F. Explosion G. Wildland Fire	A. I & C Equipment B. Criticality Equipment C. Monitor/ Analyzer D. Computer Software E. Computer Hardware	A. Freeze Protection Failure B. Seismic Qualification Deficiency C. Ventilation System/Fan D. Mechanical Equipment Failure/Damage E. Structural Deficiency/Failure F. Corrosion/Material Degradation/End of Life G. Glovebox Failure H. HEPA Filter I. Container/Package Failure

6. Radiological	7. Electrical Systems	8. OSHA Reportable/Industrial Hygiene		9. Safeguards/Security Issue
A. Clothing Contamination B. Facility/Equip/Site Contamination C. Skin Contamination D. Airborne Radiological Release E. Radiological Control Procedure Noncompliance F. External Exposure G. Intake H. Inadequate Radiological Control Job Planning I. Radiological Control Training Deficiency J. Inadequate Radiological Control Procedure K. Offsite Spread of Contamination	A. Emergency/ Backup Generator Failure B. Electrical Distribution C. Power Outage D. Electrical Wiring E. Electrical Equipment Failure F. Arc Flash	A. Electrical Shock B. Indoor Air Contamination C. Industrial Hygiene Exposure D. Injury E. Fatality F. Industrial Operations Issues (Retired) G. Industrial Equipment H. Safety Noncompliance I. Safety Equipment Failure J. Near Miss (Electrical)	K. Notice of Violation or Noncompliance from local, state, or federal agency. L. Chemical Safety M. Laser Safety N. Construction/ Demolition Safety O. Hoisting/Rigging Incident P. Forklift/Hand Truck Incident Q. Excavations/ Penetrations R. Landscaping/Mowing Beryllium Incident	A. Fitness for Duty Issue B. Material Accountability Issue C. Miscellaneous Security Issue D. Theft/Sabotage

10. Transportation	11. Other	12. EH Categories (only one per report)		13. Management Concerns	14. Quality Assurance (at least one per report)
A. Shipping Regulation Noncompliance B. Vehicle Accident/Incident C. Industrial Equipment Movement Incident D. Transportation Notice of Violation from or Noncompliance with local, state, or federal agency E. Shipping Incidents / Accidents	A. Chemical Reaction/Pressurized Drum B. Emergency Management System Failure C. Nuclear Weapons Safety Issue D. Natural Phenomena E. Suspect/Counterfeit Items F. Inadequate Design G. Subcontractor H. Procurement Deficiency/Defective Items I. Visiting Scientist/Researcher or Student Employee J. Tenants on DOE Property K. Excessed Equipment/ Material L. Supplier M. Outside Agency or Organization/ Site Visitor N. Nuclear Waste Handling Operations	A. Authorization Basis B. Conduct of Operations C. Electrical Safety D. Environmental Releases/Compliance E. Equipment Degradation/Failure F. Fire Protection and Explosive Safety G. Industrial Operations H. Injuries Requiring Medical Treatment Other Than First Aid I. Lockout/Tagout J. OS/IH K. Near Miss	L. Nuclear Criticality Safety Concern M. Radiological Control N. Radiological Skin Contaminations/ Uptakes/ Overexposures O. Safeguards and Security P. Shipping QA Q. Vehicle Accident R. Suspect/Counterfeit Items – Defective Items Z. Other than above	A. HQ Significant B. Accident Investigation – Type A (Retired) C. Accident Investigation – Type B (Retired) D. Accident Investigation – Other E. Facility Call Sheet F. Operating Experience Summary Article G. Suspect/Counterfeit Items - Defective Items Data Collection Sheet H. ARRA (Retired)	A. Program Deficiency B. Training and Qualification Deficiency C. Quality Improvement Deficiency D. Documents and Records Deficiency E. Work Process Deficiency F. Design Deficiency G. Procurement Deficiency H. Inspection and Acceptance Testing Deficiency I. Management Assessment Deficiency J. Independent Assessment Deficiency K. Safety Software QA Deficiency L. No QA Deficiency

1.4 Data Normalization

This report normalizes the data using either incidence rate (i.e., per work hours) or fire loss rate (i.e., loss per valuation) where applicable and noted. To normalize a data point using the incidence rates approach (e.g., by 200K work hours), the data point (e.g., number of reports) is divided by the total work hours of the DOE element (e.g., site) and then multiplied by 200K work hours. The resulting value represent number of reports per 200K work hours. This approach for normalizing data is also used to report OSHA incidence rates[2]. The work hours data is obtained from the DOE's CAIRS and used to calculate DOE operations incident rates.

To normalize a data point using the fire loss rate, the data point (e.g., monetary loss) is divided by the total valuation of the DOE element (e.g., site). The resulting value is reported in cents loss per $100 of valuation. This approach for normalizing data is used in the FPAS report. The valuation data is obtained from the Facility Information Management System and the Property Information Database System.

There are other approaches used to normalized data used by other organizations such as the National Fire Protection Association, Federal Emergency Management Agency, Occupational Safety and Health Administration (OSHA) and insurance industry. Examples of these normalized data approaches include reported loss per 100K population, fires per population, fire loss per 1000 dwellings, and loss per value-at-risk. Each approach is highly dependent on available data.

[2] See OSHA references: https://www.osha.gov/laws-regs/standardinterpretations/2016-08-23

2 DOE Fire Protection Losses

The DOE FPDB reports capture incident information where a fire department responded. Report information includes monetary losses, type of incident (fire and non-fire related), cause of the incident, location, and description of the incident. Sections 2.1 to 2.4 analyze and discuss trends and insights from the FPDB data and discusses main incidents that drive those trends. Section 2.5 uses machine learning text clustering to analyze the text from the description in the FPDB loss reports and identify topics that are occurring most frequently. Figure 1 illustrates the number of FPDB loss reports per year and shows that the reported records trend is flat.

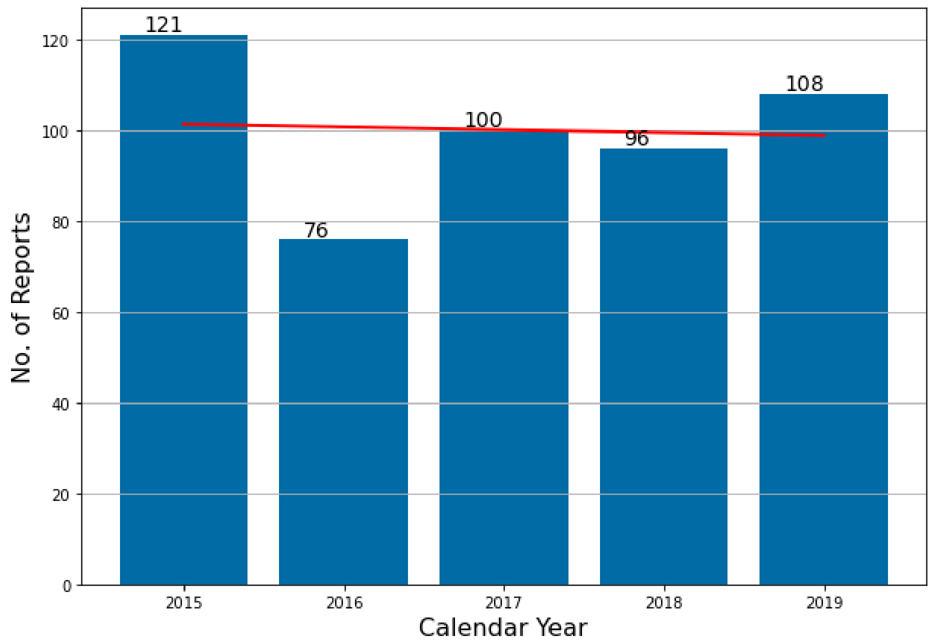

Figure 1: Number of Fire Protection Loss Reports

Figure 2 shows the number of reports normalized by 200K work hours. This figure shows that the trends for the normalized rate of FPDB loss reports are decreasing over time.

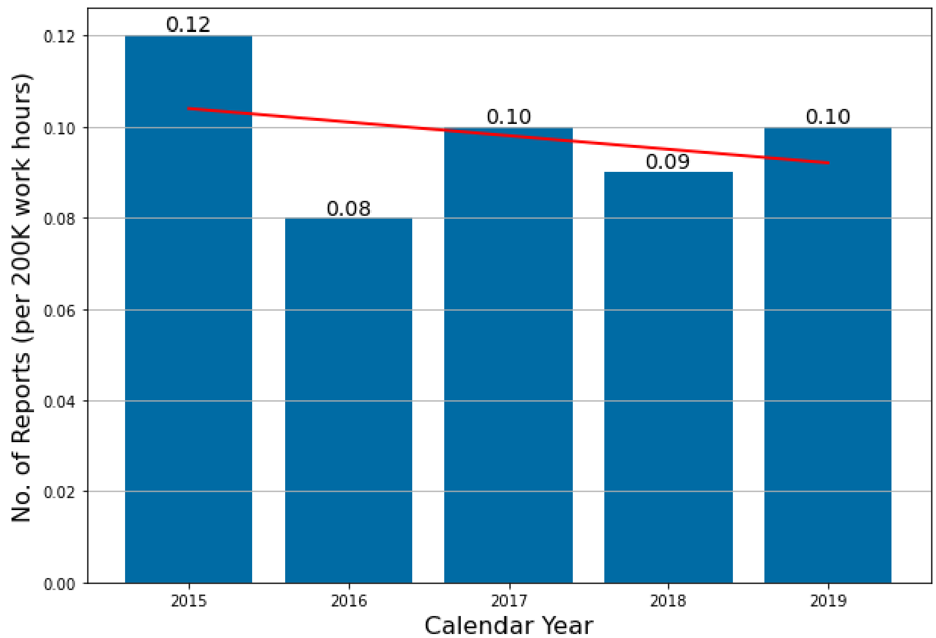

Figure 2: Number of Fire Protection Loss Reports (per 200K work hours)

2.1 Cost of Fire Protection Losses at DOE

All metrics related to the cost of fire protection losses, are shown in Figure 3 thru Figure 6. These figures illustrate that the cost related trends are increasing. Since Figure 1 show that the trend for the number of reports is flat and Figure 2 show that the number of events is slightly decreasing when normalized by work hours, this could indicate that incidents are less frequent but with higher consequence.

The trends of reported monetary losses are shown in Figure 3, i.e., fire protection losses over the five-year period and Figure 4, i.e., fire protection loss (normalized by 200K hours worked). Calendar years 2015 and 2019 show above average fire protection losses. The 2015 fire protection loss was driven by two incidents: 1) a $500K loss due to frozen pipes bursting and flooding of a building, and 2) an approximately $400K equipment loss due to a transformer fire. These two incidents resulted in more than half of the 2015 total monetary losses. Increases in 2019 were also driven by two events, one involving a gas chromatograph ($750K loss), and another involving a lightning-initiated fire ($700K loss). These two events accounted for more than half the total loss in 2019.

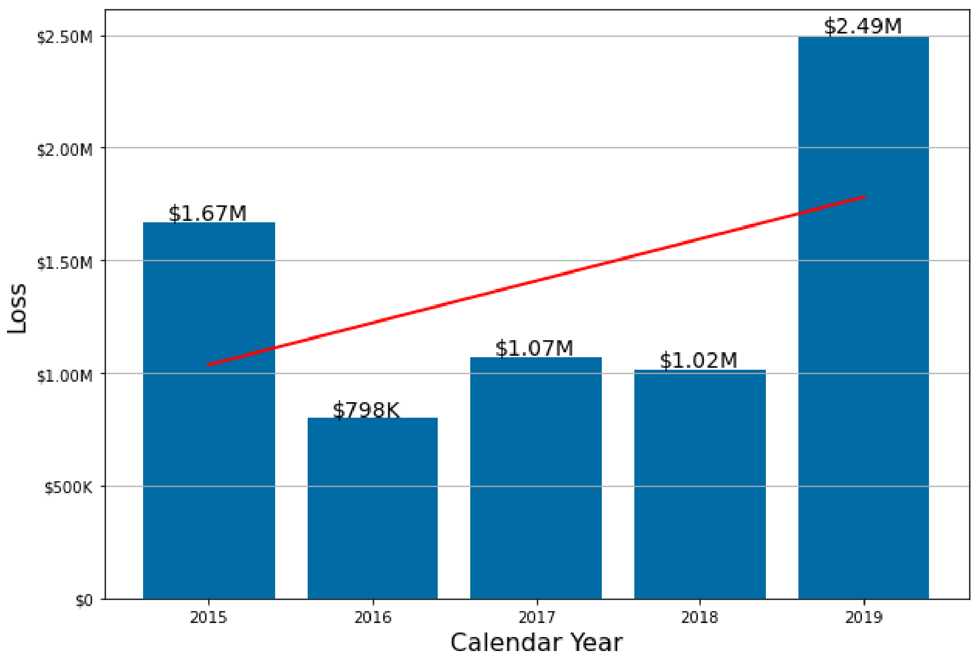

Figure 3: Fire Protection Losses per Calendar Year

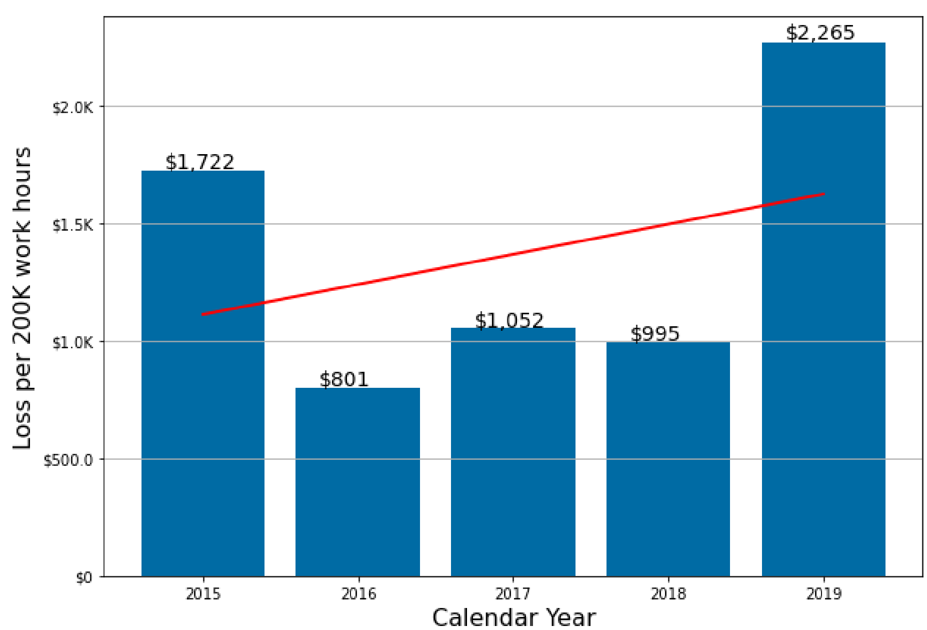

Figure 4: Fire Protection Losses in Dollars Rate by Hours Worked at DOE

Figure 5 show the DOE fire loss in dollars rate by work hours and its trend. The figure shows an increasing trend over the reporting period.

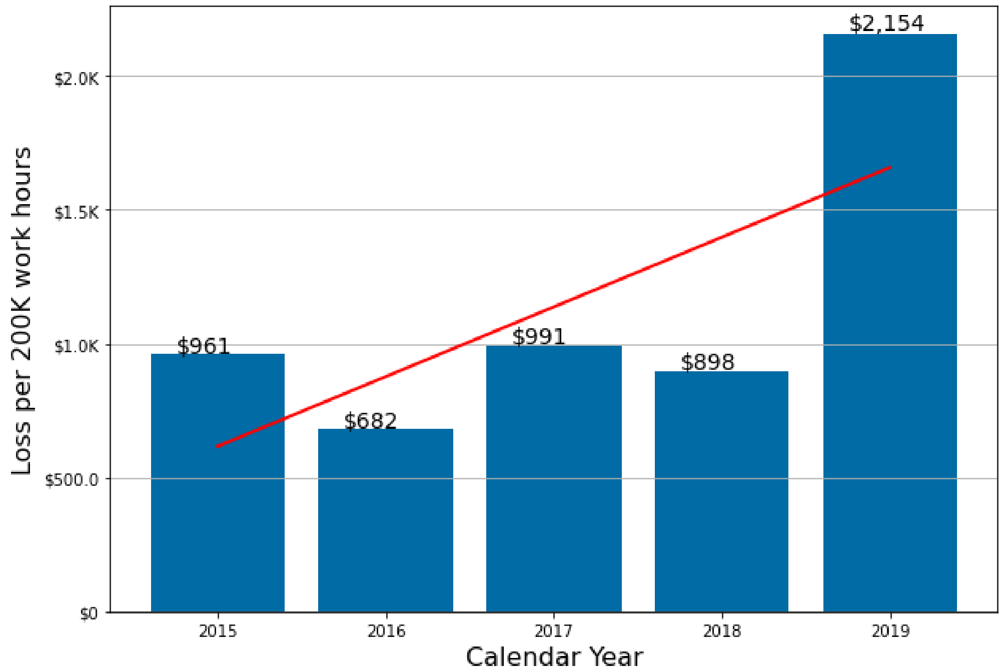

Figure 5: Fire Losses in Dollars Rate by Hours Worked at DOE

Figure 6 shows the trend of the DOE fire loss rates (cents per $100 valuation). This figure shows an increasing trend of the fire loss rate over the reporting period. This data was obtained from the 2019 DOE Fire Protection Annual Summary.

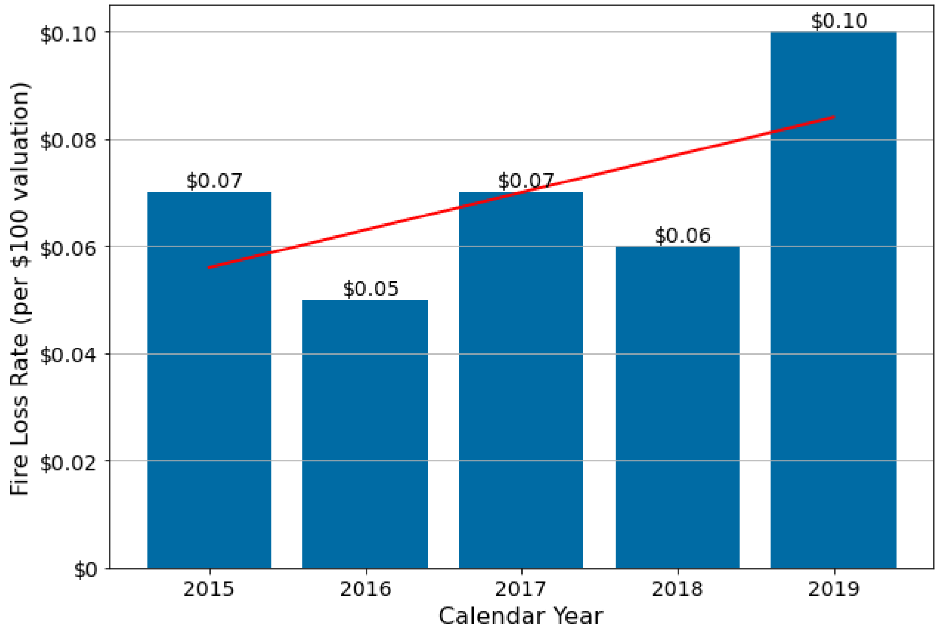

Figure 6: DOE Fire Loss Rates 2015-2019 (cents per $100 valuation)

2.2 Types of Reported Fire Protection Losses

This section discusses the trends of fire protection losses by loss type. Entries in the FPDB are classified as either fire or non-fire loss reports. Fire-related reports include those resulting from fire/smoke losses (e.g., building, brush, vehicle, and other). Non-fire reports include those where damage was sustained as a consequence of non-fire events involving fire protection systems (including leaks, spills, and inadvertent releases) such as:

- Weather related (e.g., frozen sprinkler pipe damage)
- Electrical malfunction (not classified as fire)
- Transportation (cargo and vehicle accident) that resulted in an inadvertent fire protection system actuation or
- Mechanical malfunction of a fire protection system

2.2.1 Reported Fire Losses

Figure 7 shows the number of fire loss reports and their type within the FPDB. Figure 8 shows that the trend for fire monetary losses (e.g., fire/smoke) related events is increasing. This is driven by the two 2019 events which added up to more than half the total fire loss of 2019. Figure 9 shows the distribution of the types of fires. This figure shows that the "building" and "vehicle" monetary losses are mostly constant over the period. The "other" category losses were small in 2017 and 2018. The highest fire loss was in 2019. The "brush" losses have been constantly increasing during the last three years.

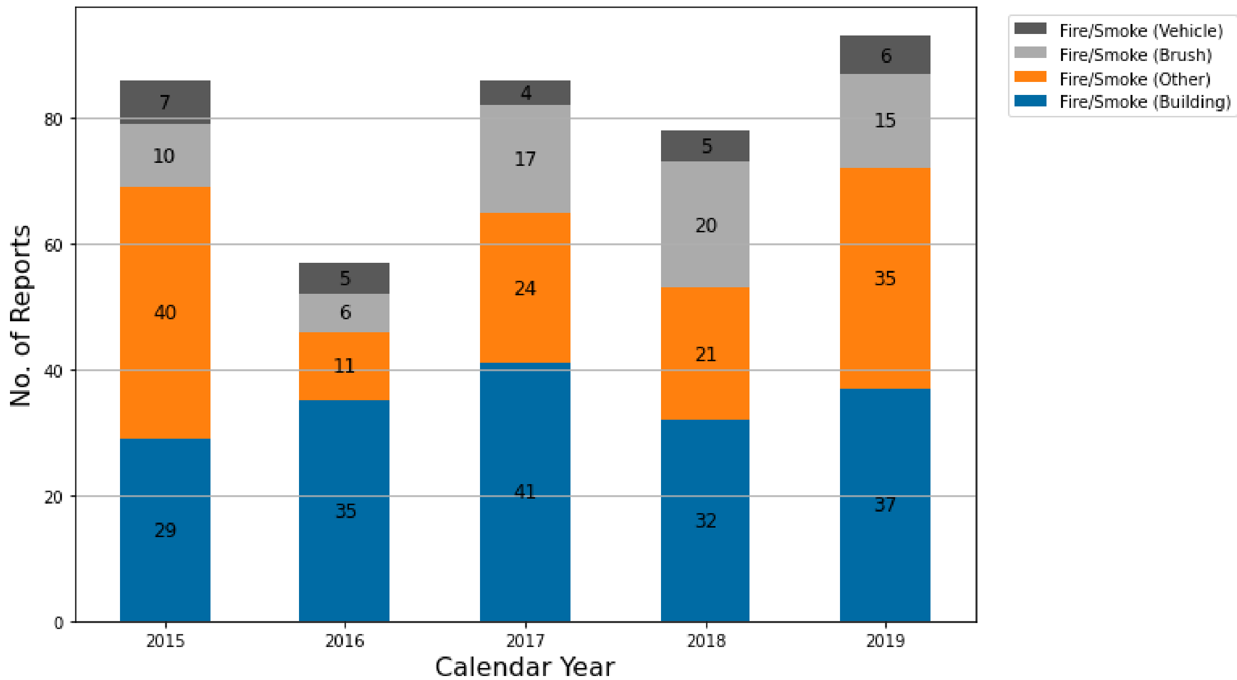

Figure 7: Number of Fire Loss Reports by Type

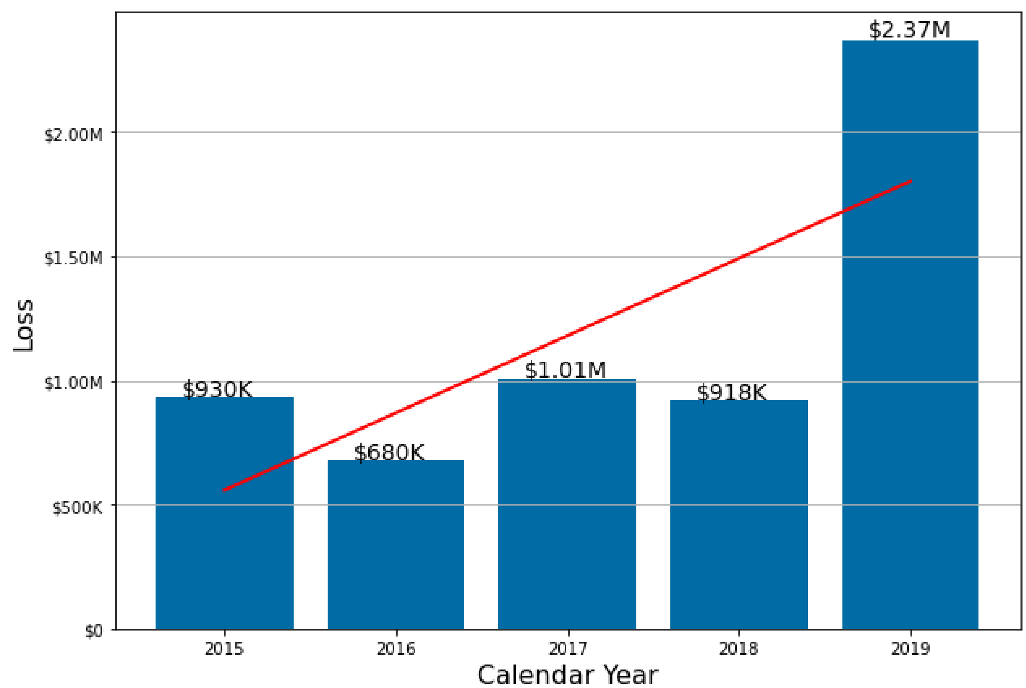

Figure 8: Fire Losses

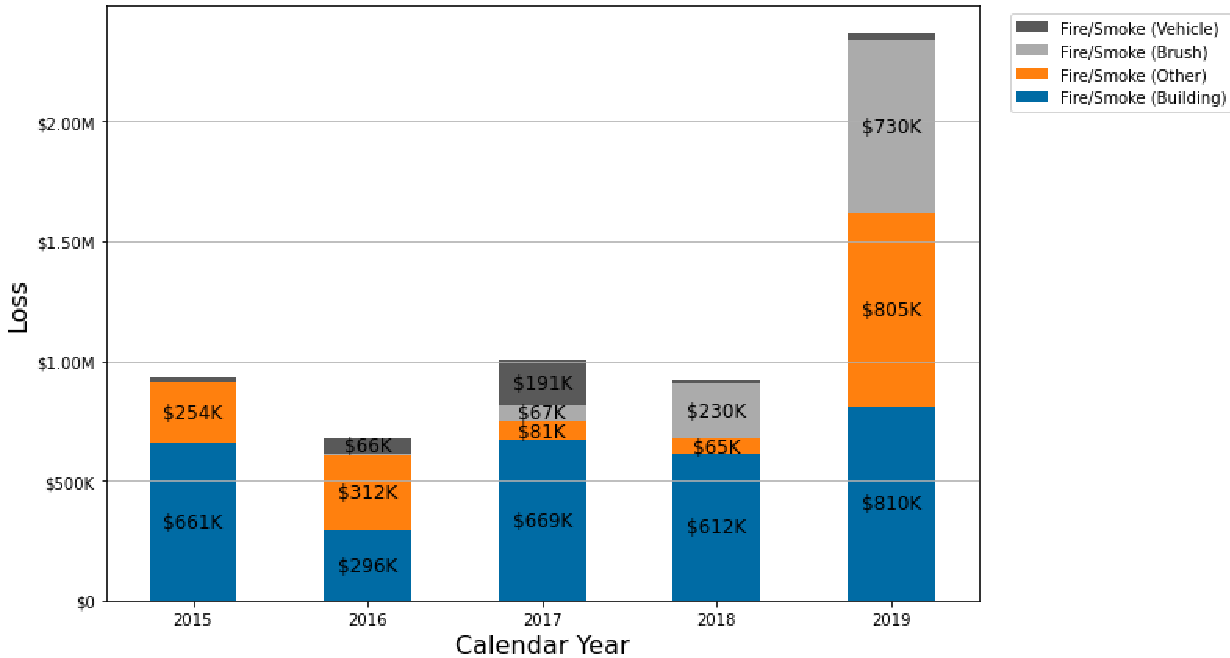

Figure 9: Breakdown of Fire Losses by Type (values of $50K or more shown)

2.2.2 Reported Non-Fire Losses

Figure 10 shows the number of non-fire reports at DOE. Figure 11 shows the total reported non-fire losses at DOE. The figure shows that the trends are decreasing. In 2015, two reports resulted in $633K losses, otherwise the yearly non-fire losses are around $100K per year. The figures show a decreasing trend showing a sharp drop of non-fire reports and monetary losses in the years after 2015.

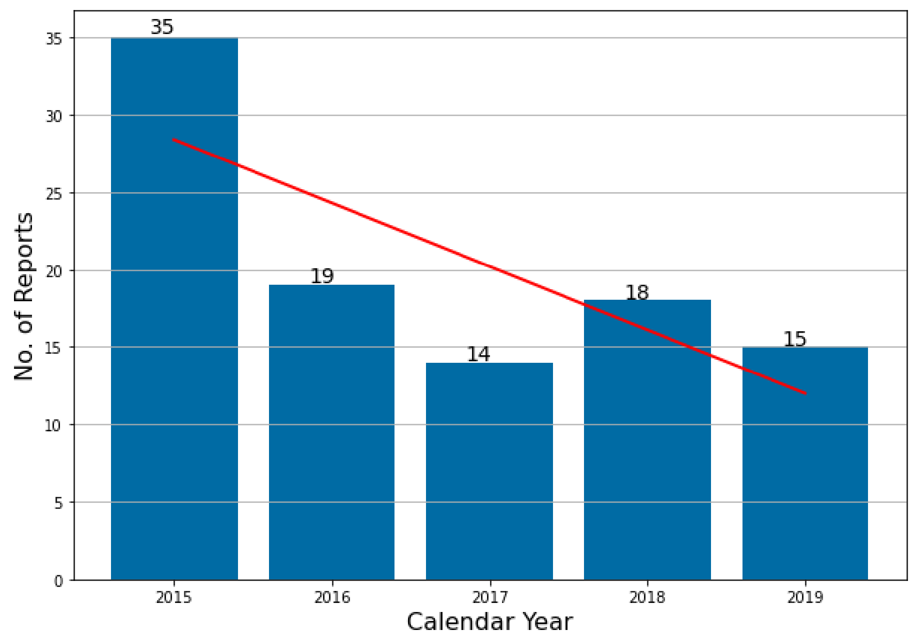

Figure 10: Number of non-Fire Loss Reports

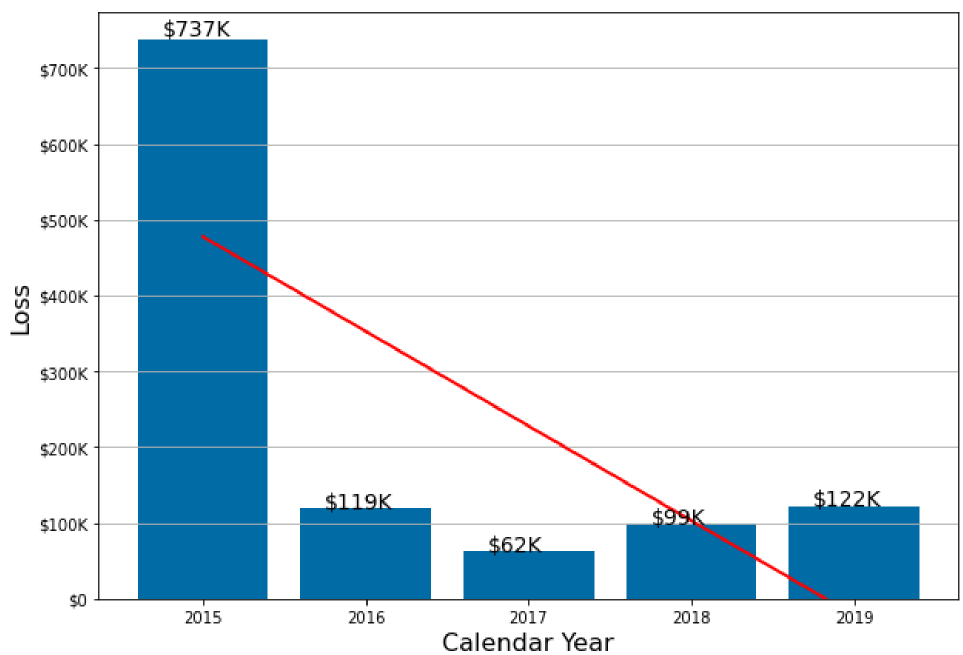

Figure 11: Non-Fire Losses

2.2.3 Types of Reported Fire Protection Losses Trends

This section calculates and discusses the trends for each type of fire protection report in the 2015 to 2019 time period. Figure 12 shows the number of reports for each type of fire.

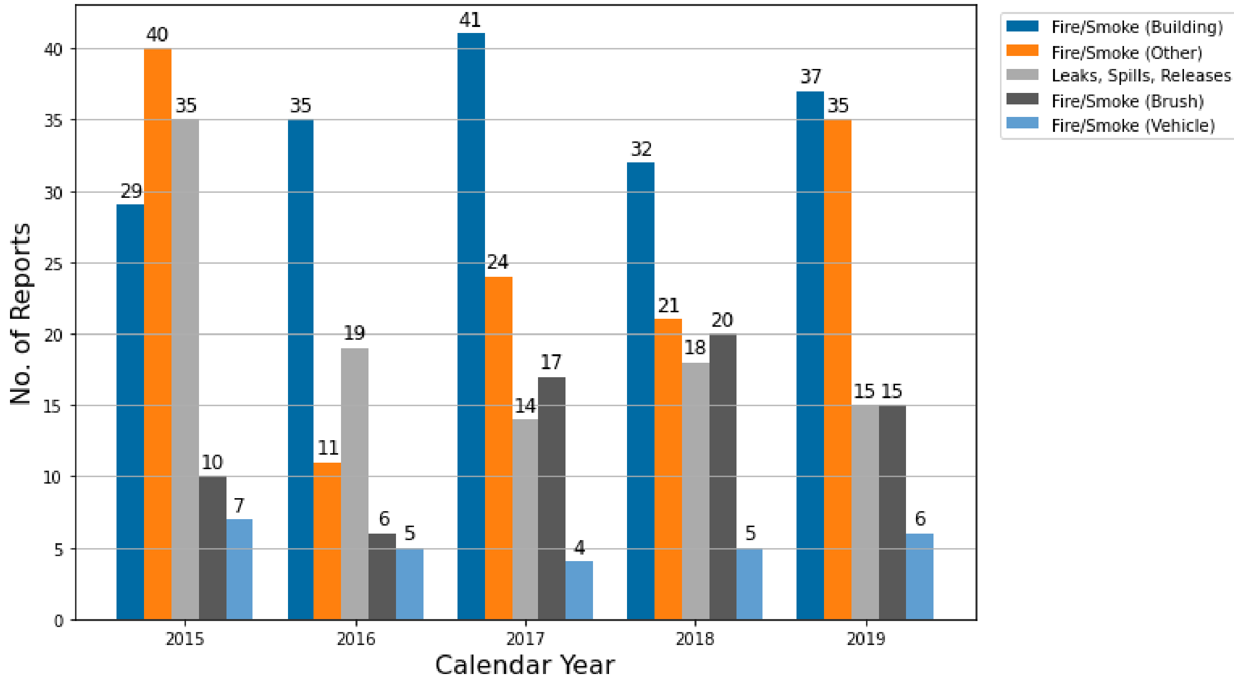

Figure 12: Fire Protection Losses by Type

Table 2 shows the slope of the trend for the number of reports and the monetary losses per year for each fire protection loss type. The slope of the trend shows the degree of increase (or decrease) and is calculated for both number of reports per calendar year and monetary losses per calendar year for each type of report. A positive slope of the trend indicates that the linear trend is increasing by the specified amount per year. For example, the "Fire/Smoke (Brush)" slope for number of reports is 2.4 and the losses is $168,324. This means that on average the number of reports are increasing by 2.4 reports per year and the loss is increasing by $168,324 per year. Note that these slope values are the values of the linear trends slope shown in Figure 13 and Figure 14. A negative slope of the trend indicates that the linear trend is decreasing the specified amount per year. This table is used to determine which are the highest increasing trends per type of report (or cause of report in Section 2.3). Table 2 shows that the Fire/Smoke (Brush) type of reports is increasing the most in any of the trends calculated. The Fire/Smoke (Building) and Fire/Smoke (Other) had the second and third highest increasing trends for the number of reports. When evaluating the trend of monetary losses, the trend of the Fire/Smoke (Other) and Fire/Smoke (Building) have the second and third highest increasing trends respectively. Leaks, Spills, Releases and Fire/Smoke (Vehicle) have both decreasing (negative) trends. Only the highest increasing trend for number of reports and for monetary losses are discussed in more detail.

Table 2: Trends of the Fire Protection Loss Types

	Slope of the Trend				
	Fire/Smoke (Other)	**Leaks, Spills, Releases**	**Fire/Smoke (Brush)**	**Fire/Smoke (Building)**	**Fire/Smoke (Vehicle)**
Number of Reports per year	0.0	-4.1	2.4	1.3	-0.2
Monetary Loss per year	$85,332	-$124,923	$168,324	$61,567	-$4,199

Fire/Smoke (Brush) had the highest increasing trend (i.e., highest increasing slope) by number of reports and by monetary losses. Figure 13 shows the trends of the number of reports. Figure 14 shows the trends of the monetary loss. The increasing trend in monetary losses are driven mainly by three reports. A 2017 $50K loss due to an equipment failure on a transmission line which resulted in a vegetation fire of over two acres. A 2018 $200K loss of a wildland fire caused by a high voltage power supply line. A 2019 $700K loss due to a lightning induced wildland fire which involved over 112K acres.

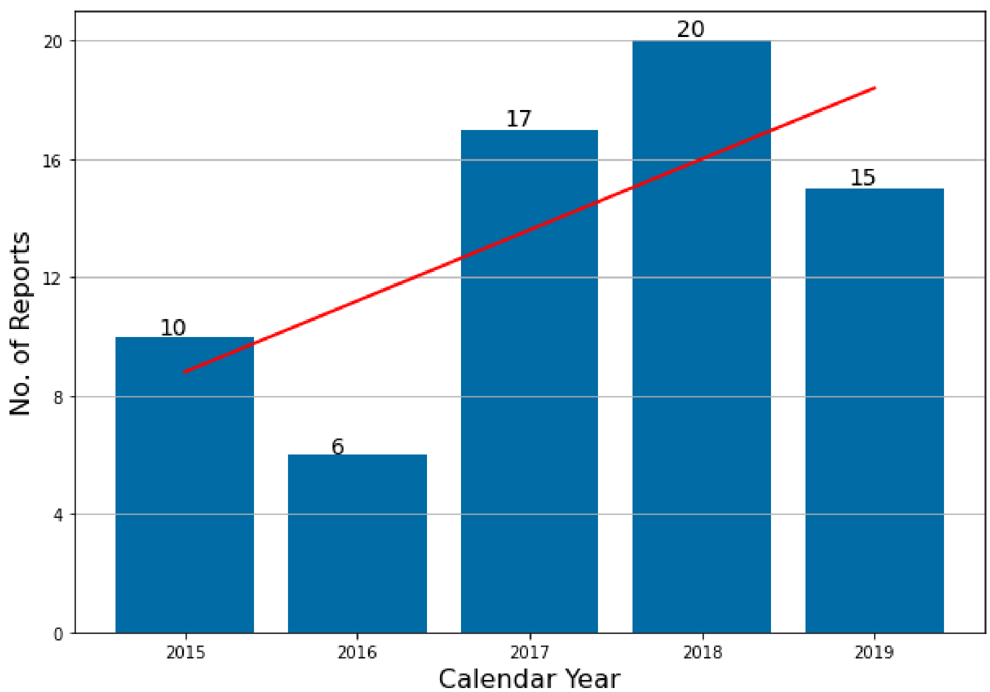

Figure 13: Number of Fire/Smoke (Brush) Loss Reports

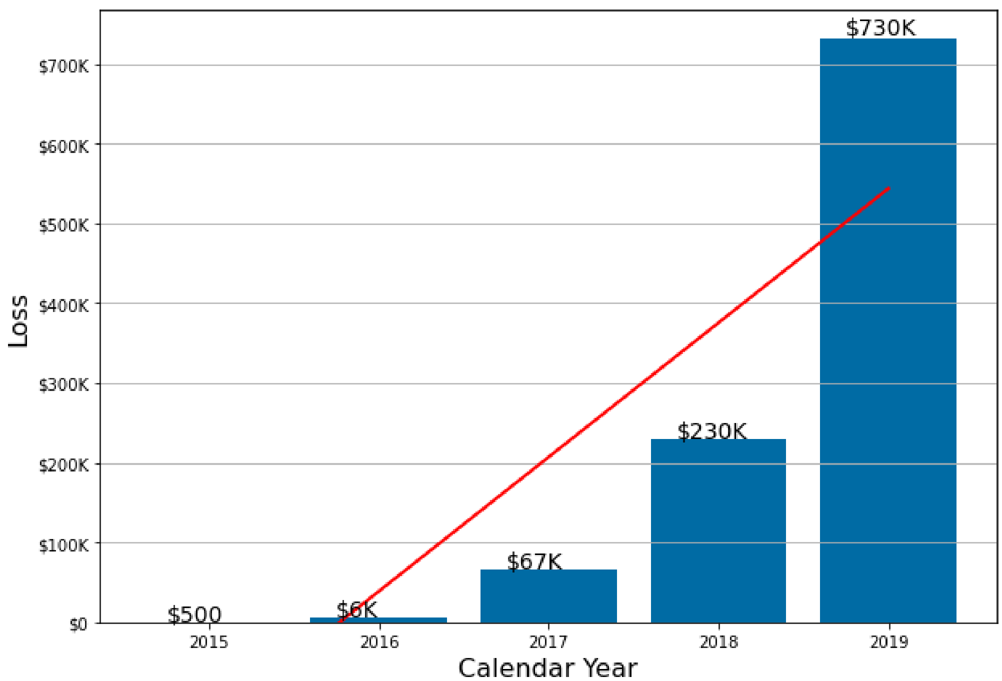

Figure 14: Fire/Smoke (Brush) Losses

2.3 Causes of Reported Fire Protection Losses

The fire protection data includes classification of each report by the identified cause. These causes include employee related, electrical, design/material related, weather related, procedure related, other and unspecified. Figure 15 shows the number of reports by identified cause in a stacked bar plot. Figure 16 shows the number of reports by identified cause in unstacked bar plot.

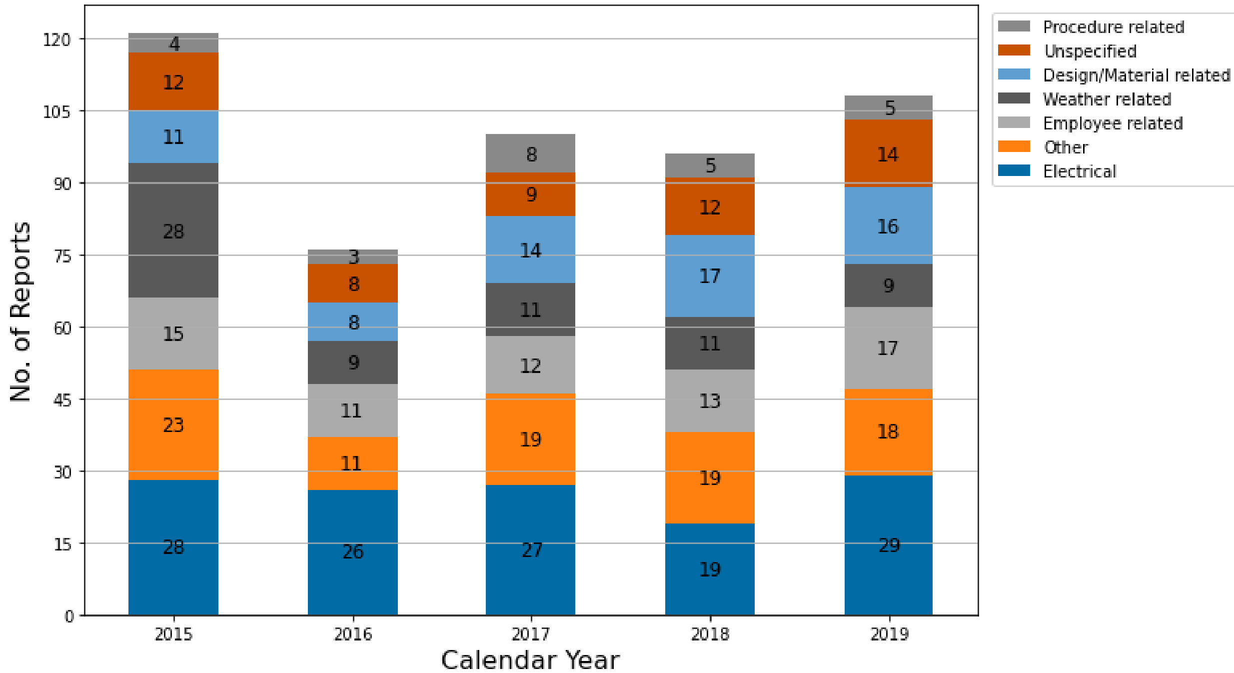

Figure 15: Number of Fire Protection Reports by Cause

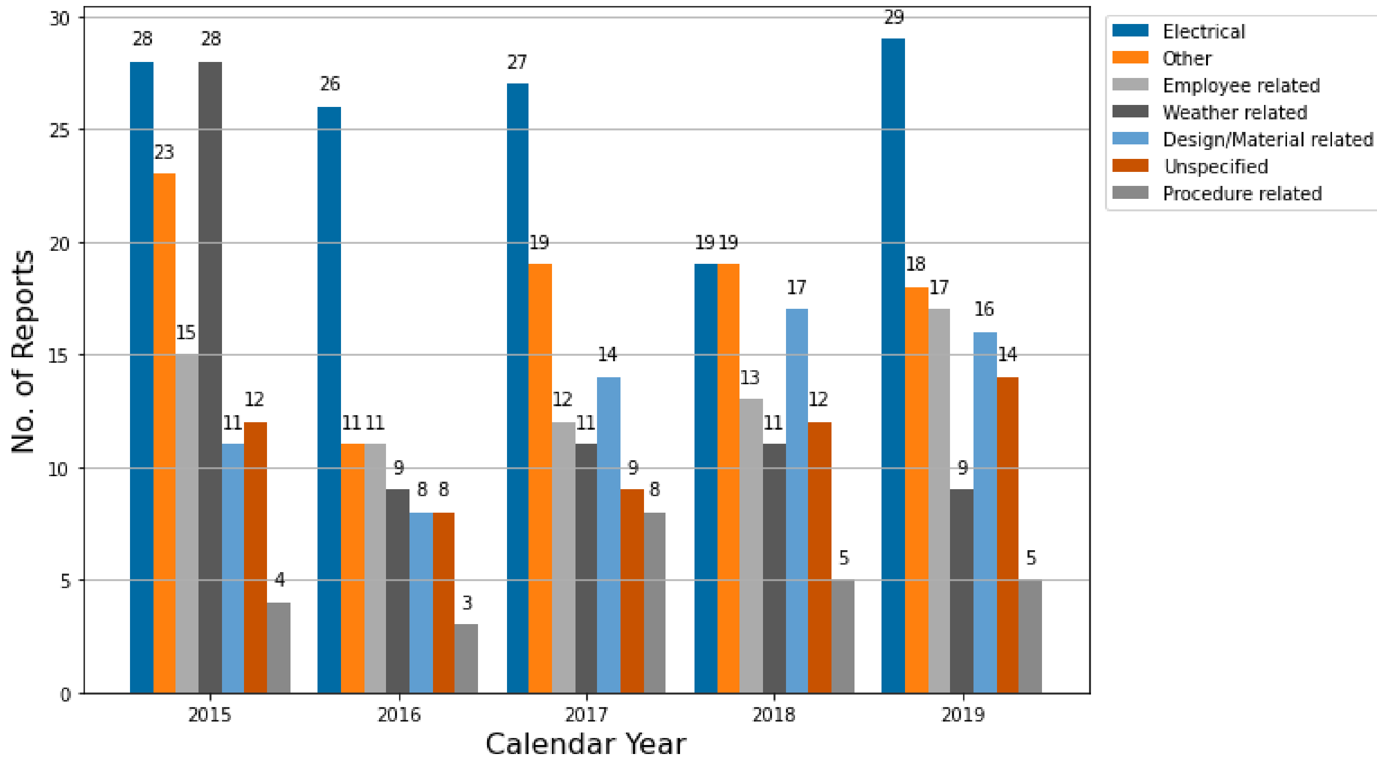

Figure 16: Number of Fire Protection Reports by Cause

The Figure 17 shows the monetary loss by cause. The biggest losses are in 2015 related to weather and electrical causes and in 2019 related to weather and other causes. The 2015 weather related highest losses were due to two reports where frozen pipes burst resulting in flood losses of $500K and $133K. The 2015 electrical related causes are due to a heating ventilation and air conditioning related fire that resulted in a $398K loss and a transformer fire that resulted in a $434K loss. The 2019 weather related cause was due to the Idaho National Laboratory fire that resulted in a $700K loss and the “other” cause mainly due to a gas chromatograph fire that resulted in $750K loss.

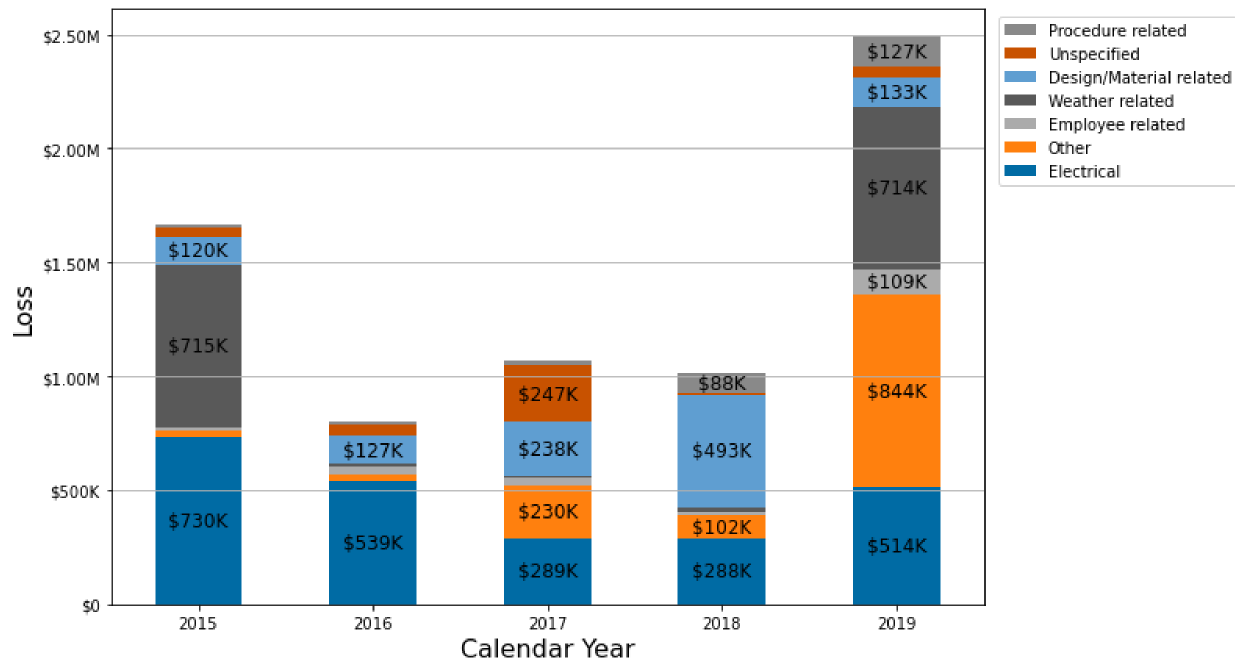

Figure 17: Fire Protection Losses by Cause (values of $50K or more shown)

Table 3 shows the trend slope of the number of reports and the trend slope of the monetary losses per year for each cause. This table shows that the Design/Material related cause has the highest increasing trend by number of reports and the second highest increasing trend for monetary losses. The "other" cause had the highest increasing monetary loss trend. The "Other" cause is further discussed in Section 2.3.2. Discussion on interpretation of the slope of the trends is provided in Section 2.2.3. Only the highest increasing trend for number of reports and for monetary losses are discussed in more detail.

Table 3: Slope of the Trend of the Cause of Reports over the period between 2015 to 2019

	Slope of the Trend						
	Other Related	Employee Related	Electrical Related	Unspecified Related	Design/Material Related	Weather Related	Procedure Related
Number of Records per year	-0.2	0.6	-0.5	0.8	1.9	-3.6	0.4
Monetary Loss per year	$170,171	$16,909	-$68,305	-$2,177	$39,150	$306	$30,045

Figure 18 and Figure 19 show the plots for the number of reports and monetary losses for the Design/Material related causes. Figure 18 shows the number of reports has been increasing year over year. Figure 19 shows larger than averages fire protection losses in 2017 and 2018. In 2017 a natural uranium metal material fire resulted in a $200K loss, while in 2018 a loss of $362K was reported due to investigation support costs for a facility that was being transitioned from contractor to DOE owned.

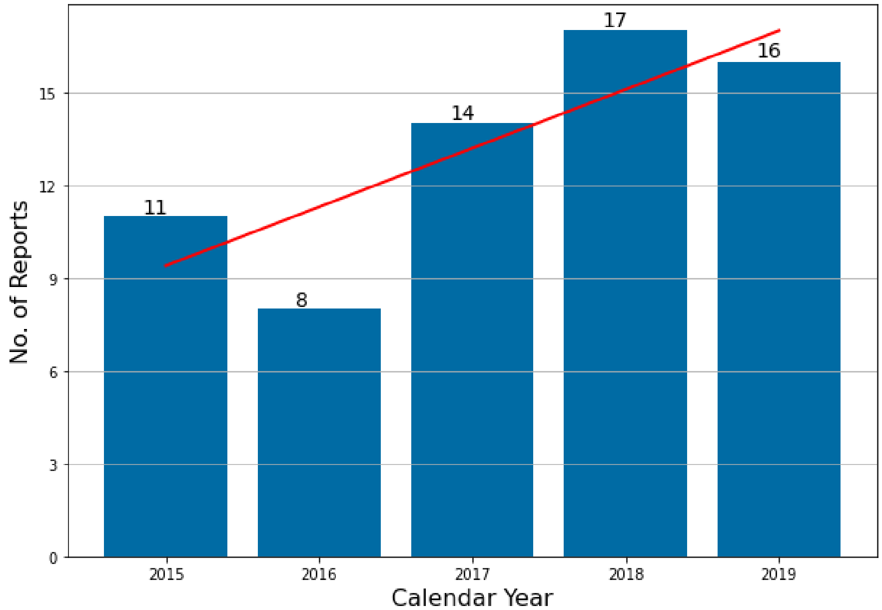

Figure 18: Number Fire Protection Reports related to Design/Material Cause

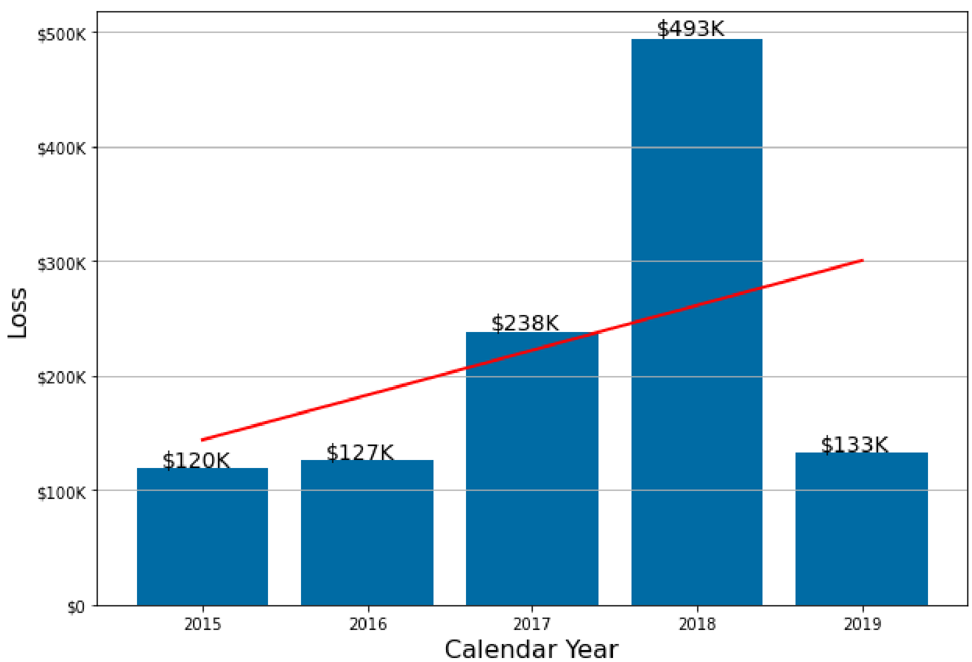

Figure 19: Fire Protection Losses related to Design/Material Cause

2.3.1 Employee-Related Causes

Table 3 shows that the employee-related cause had the third highest increasing trend in both the number of reports and monetary losses. The trend of the number of reports of the employee-related causes is shown in Figure 20. Figure 21 shows the trend for the monetary losses for the same cause. Both figures show that the trends are increasing. Figure 21 shows an above average monetary loss in the calendar year 2019. The two largest losses ($57K and $40K) occurred in 2019. One was related to an analytical microwave where an employee operating the equipment used a wrong probe. The other was related to a vehicle accident where an employee experienced a medical condition, crashed into a building, shearing the fire riser off at the base and flooding the building. The normalized monetary losses do not show any insights different from the non-normalized.

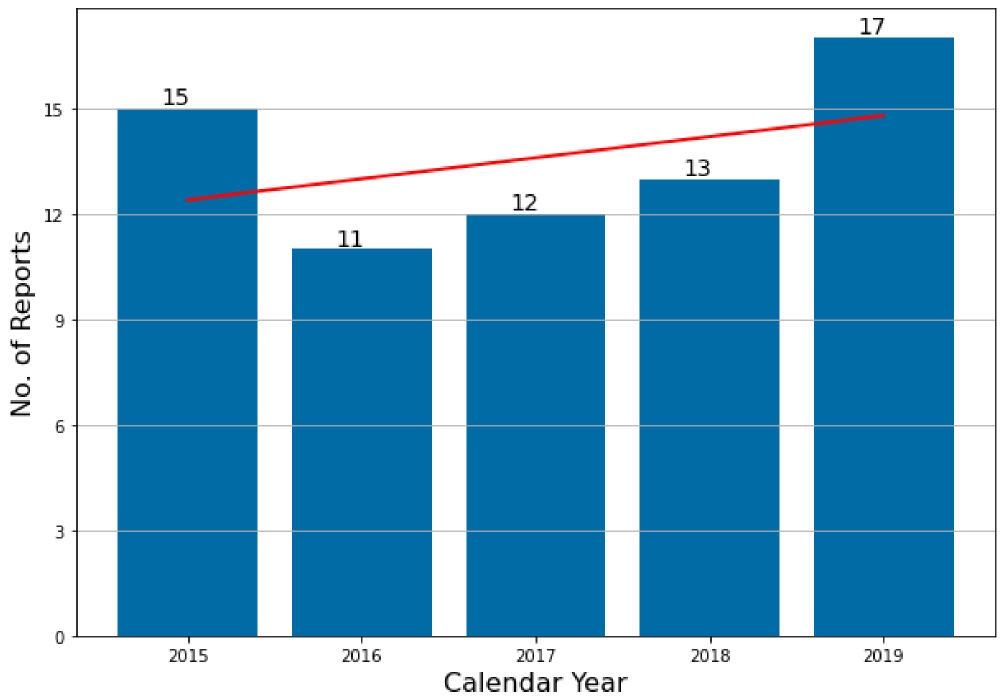

Figure 20: Number of Fire Protection Reports related to Employee Cause

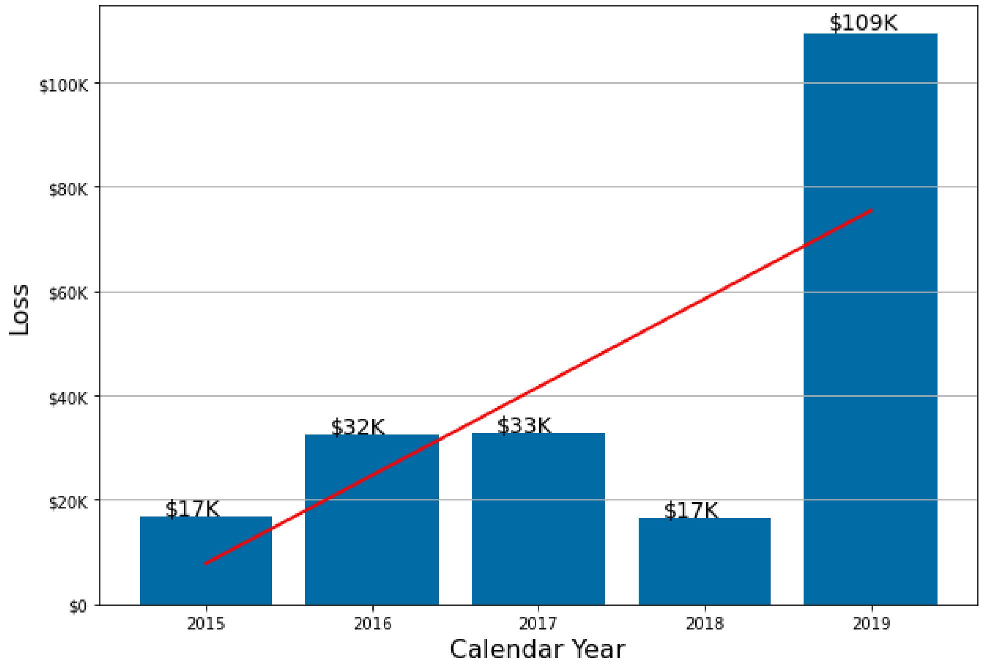

Figure 21: Fire Protection Losses Caused by Employees

2.3.2 Other Causes

Table 3 showed that the Other Related cause had the fourth highest increasing trend for the number of reports and highest increasing trend for the monetary losses. The trend of the number of the Other Related cause reports is shown in Figure 22 and the trend for the monetary losses in Figure 23. Figure 23 shows a large increasing trend. This is caused by the $750K loss caused in the 2019 gas chromatograph fire. Other years had losses of less than $250K. The normalized monetary losses do not show any insights different from the non-normalized.

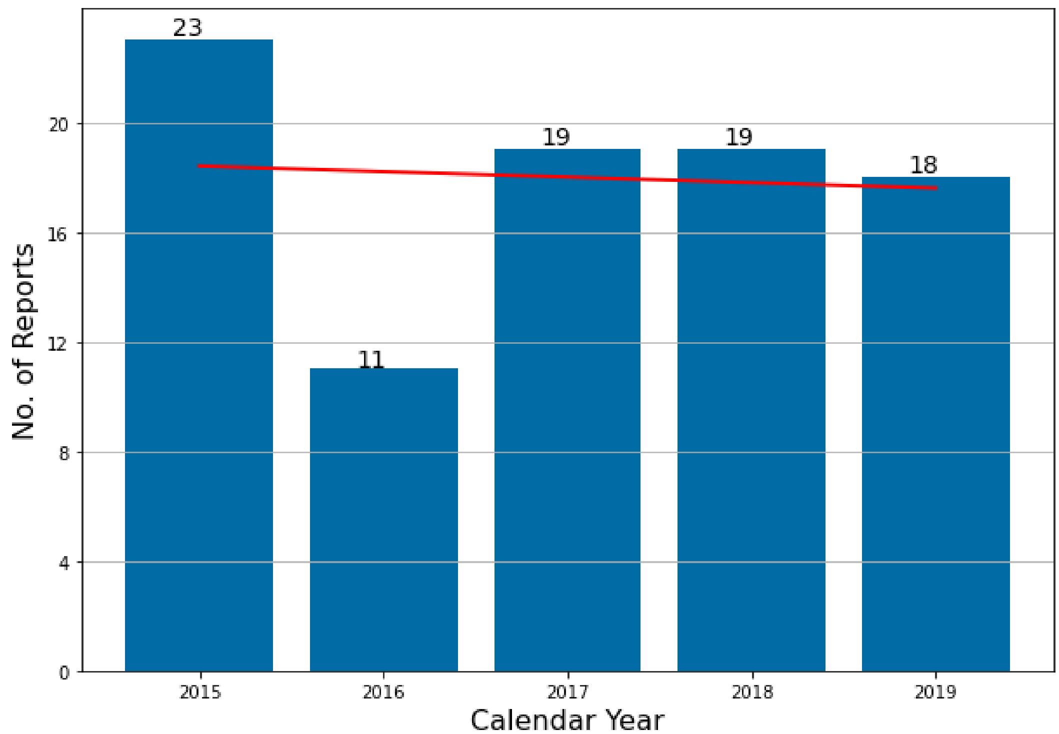

Figure 22: Number of Fire Protection Reports related to Other Cause

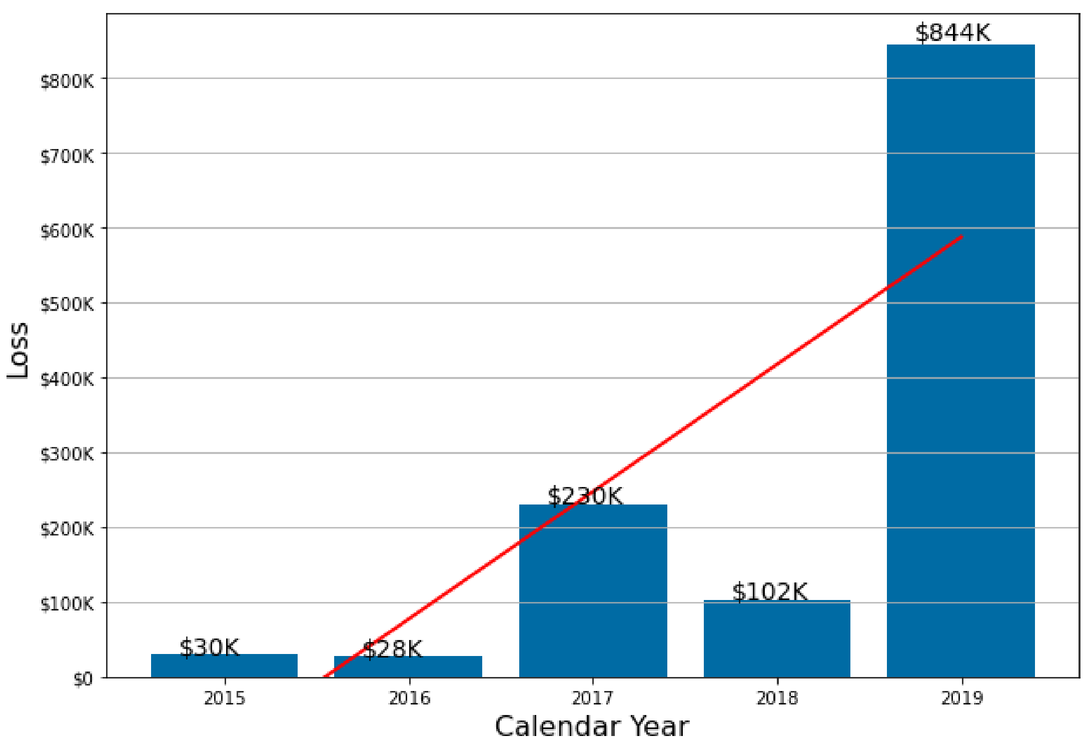

Figure 23: Fire Protection Losses related to Other Cause

2.3.3 Unspecified Causes

Table 3 showed that the "unspecified" related cause had the second highest increasing trend for the number of FPDB reports and a slight decreasing trend (second to last increasing trend) for the monetary losses. The trend of the number of the "unspecified" cause related reports is shown in Figure 24 and the trend for the monetary losses in Figure 25. Figure 25 shows an above average loss in 2017 which was the result of an incident where a shed, housing two generators, was destroyed by a fire for a total loss of $235K. The normalized monetary losses do not show any insights different from the non-normalized.

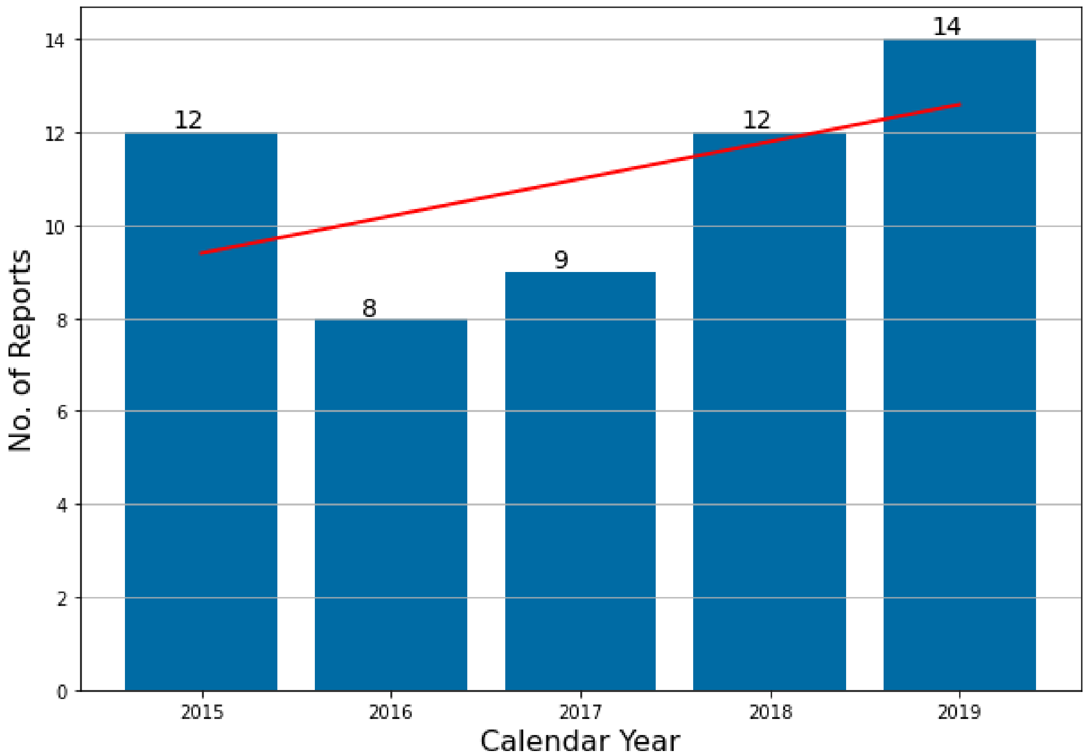

Figure 24: Number of Fire Protection Reports related to Unspecified Cause

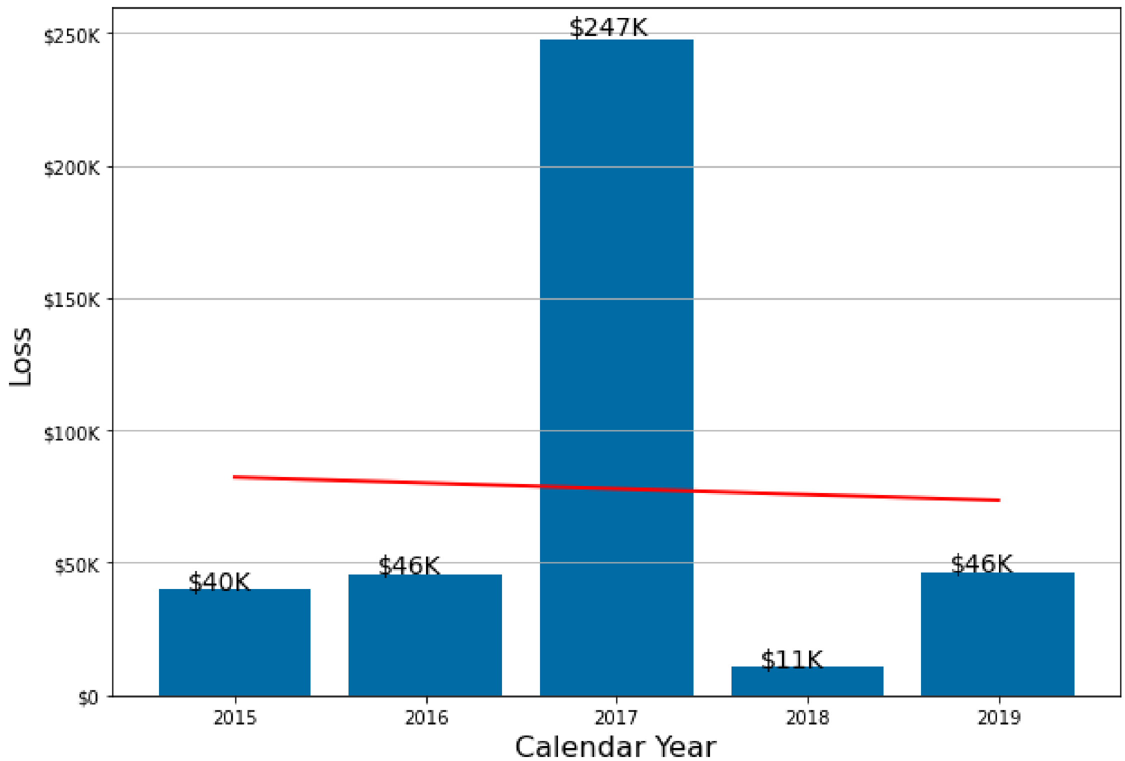

Figure 25: Fire Protection Losses related to Unspecified Cause

2.4 DOE Fire Protection Losses Greater than $50K

This section discusses the trends for fire protection losses that resulted in a monetary loss greater than $50K. The Section 2.6 and 2.7 discuss the text and topic analysis of the reports. Figure 26 shows that the number of reports for fire protection losses greater than $50K is slightly decreasing. Figure 27 shows that the trend for fire protection losses greater than $50K is increasing. This could indicate that incidents are becoming less frequent but resulting in higher monetary losses.

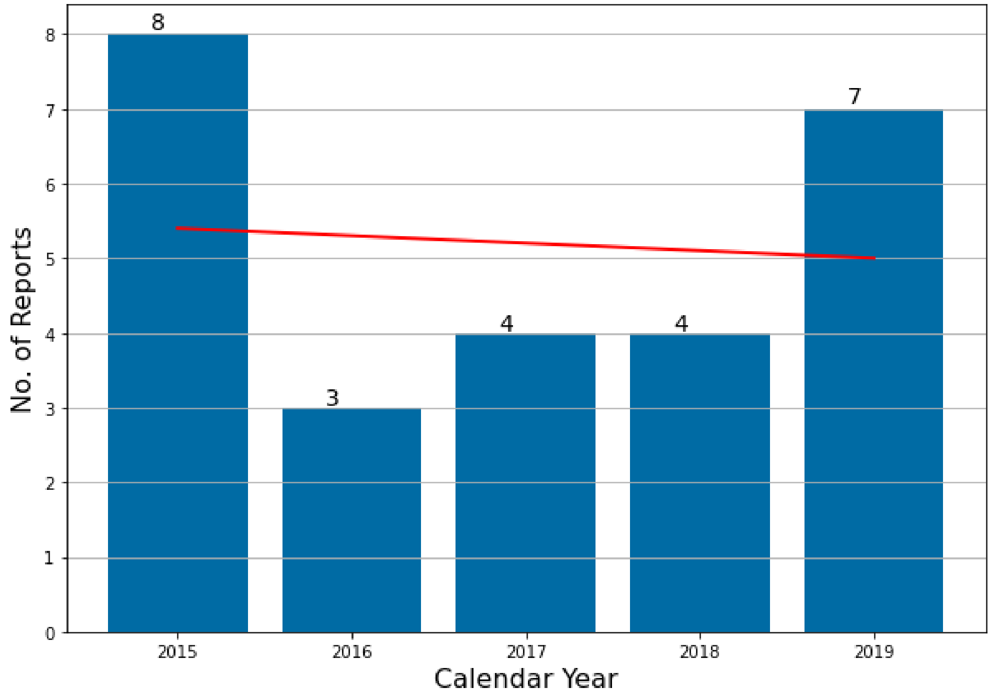

Figure 26: Number of Fire Protection Reports with Losses Greater than $50K

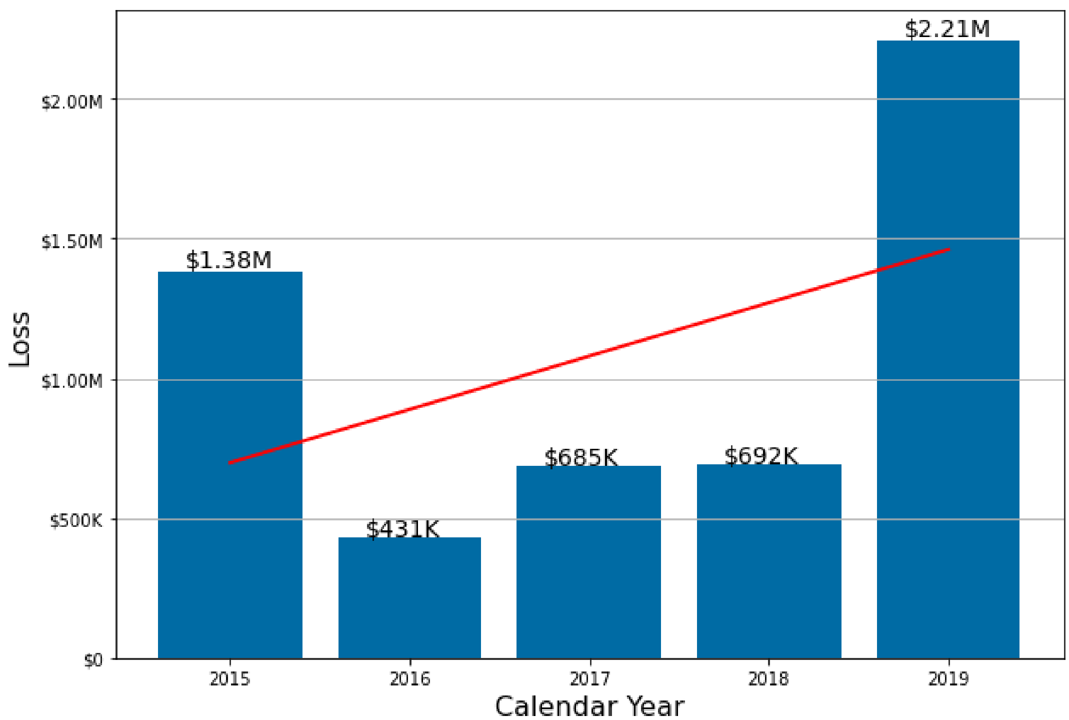

Figure 27: Fire Protection Losses Greater than 50K

Figure 28 show the types of the fire reports that resulted in fire protection losses greater than $50K. This figure shows that the fire/smoke (building) type resulted in the most events year over year and the fire/smoke (other) type resulted in at least one report in three separate years, three in 2015, one in 2016 and one in 2019.

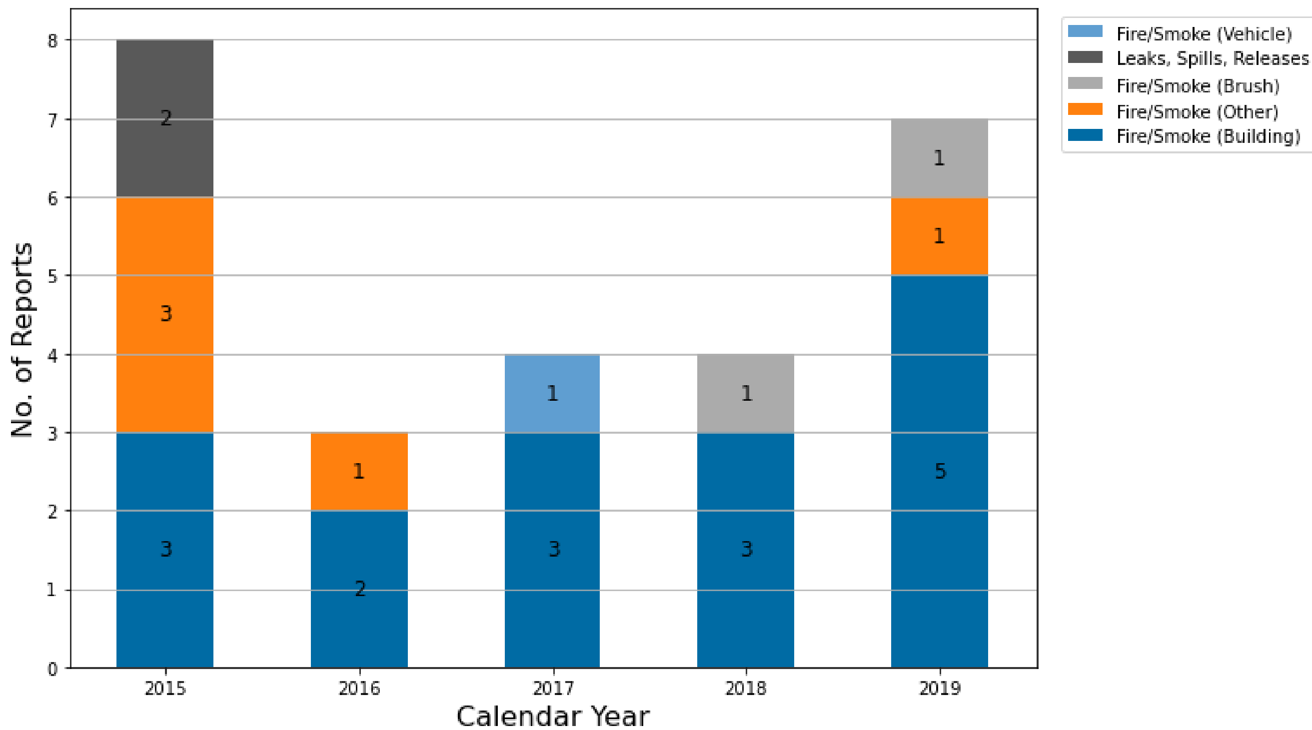

Figure 28: Number of Fire Protection Reports with Losses Greater than $50K by Type

Figure 29 shows the causes of the fire reports that resulted in fire protection losses greater than $50K. This figure shows that the electrical and design/material causes resulted in at least one report in each calendar year. The other causes only had two or less reports in only one or two years.

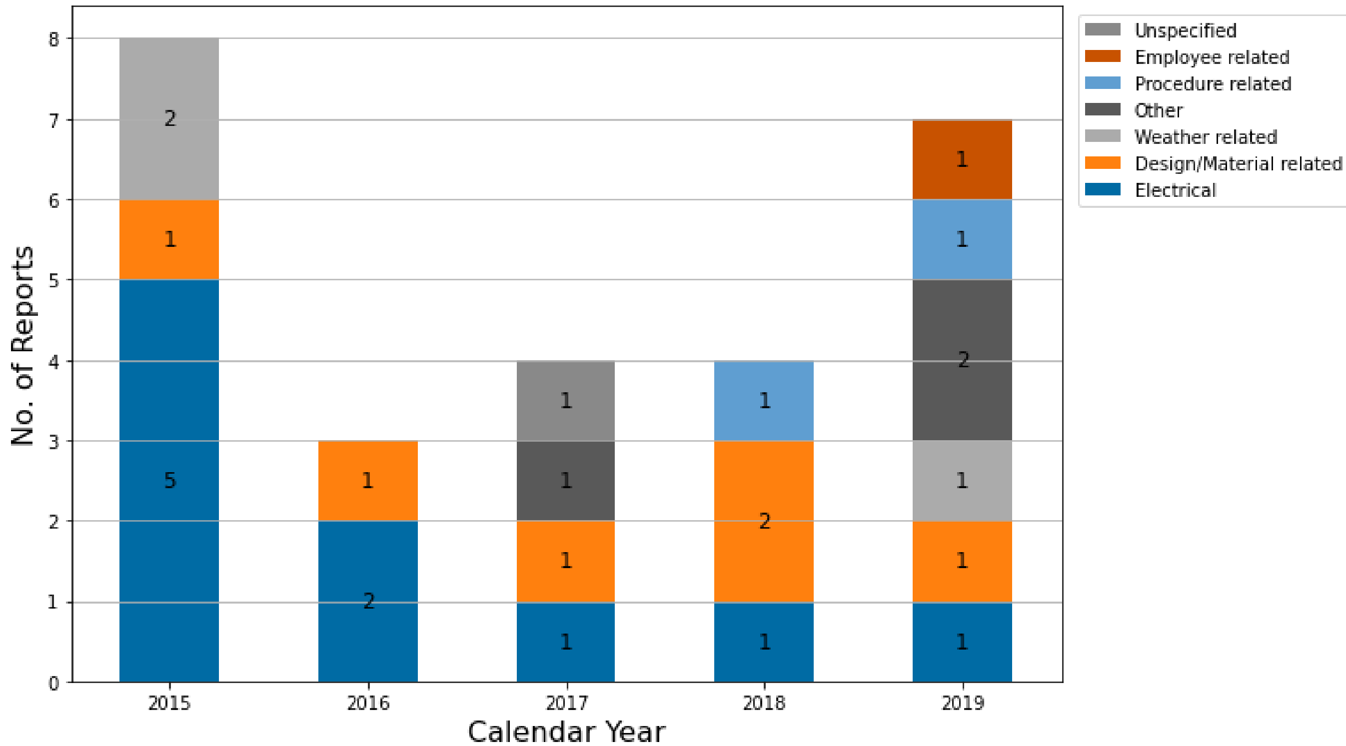

Figure 29: Number of Fire Protection Reports with Losses Greater than $50K by Causes

2.5 Fire Protection Loss Reports Machine Learning Text Analysis

2.5.1 Introduction to Machine Learning Text Analysis

DOE's Office of ES&H Reporting and Analysis has been exploring applications of machine learning and developed tools to analyze safety report data. The approaches shown here can be used to aid ES&H practitioners in the identification of high frequency and determine the relative importance of terms of interest (i.e., recurrent words that could imply a systemic issues), confirm current terms of interest, and quickly analyze and provide insights on text data. The first objective of this section is to provide some background on the current machine learning algorithms, text analysis approaches utilized and explain how natural language processing was used to obtain insights, which are discussed below. The second objective is to provide the results of the recurrent terms that were obtained from the text analysis. Readers only interested in the results should proceed to Section 2.5.2.

Machine learning methods are used in this analysis to cluster[3] fire protection loss reports and identify terms that occur within a cluster of reports. To accomplish this, all the fire protection loss report descriptions (i.e., unstructured data[4]) between the trended years (i.e., 2015-2019) were fed into an unsupervised machine learning algorithm[5] which clusters similar reports together and identifies the cluster terms. The common terms (unigrams and bigrams) are scored using term frequency-inverse document frequency (TF-IDF)[6] which can then be used to identify topical areas and reports where common or recurrent issues might exist and could then be addressed by decision makers or safety personnel.

The main purpose of the text analysis is to use machine learning to review the 500+ fire protection loss reports that occurred between 2015-2019 and identify recurrent terms. The unsupervised machine learning algorithm can review the reports and identify the top terms in a short amount of time. An ES&H subject matter expert can then review the terms and identify those that resulted in high average scores and justify further evaluating the if these terms or a cluster of reports represent an issue that could be addressed. This saves resources, improves consistency, and removes personal bias from the analysis. However, bias may be introduced when using machine learning algorithms through training data, automation of data cleaning, algorithms, and use of custom stop-words. Bias can skew the data as well as impact the results.

Machine learning is often referred to as a black box where it is difficult to describe and understand how an algorithm deducted its result from its input data. The authors have attempted to be as explicit as possible in discussing how the results are developed and how they are used to obtain insights. This analysis used unsupervised machine learning algorithms that are not affected by training data bias or affected by the input data structure (e.g., predefined types and causes of fires). The text analysis uses a list of "stop-words" to remove words that appear too often in a dataset to provide meaningful insights.

---

[3] The clustering technique used is DBSCAN with PCA and TSNE dimensionality reduction. The technique is used to identify the clusters with the largest number of reports.

[4] Unstructured data is free form data that has not been structured in a predetermined way (e.g., text description, survey responses, video, audio, images, etc.). Structured data has a predetermined form which can include text or numerical (e.g., categorical value, dates, years, quarters, quantitative values such as age, monetary values, etc.).

[5] Machine learning technique where the algorithm allows the model to discover patterns without training data (i.e., feeding the algorithm what the result looks like)

[6] Term frequency-inverse document frequency (TF-IDF): statistical measure that evaluates how important a word is to a document in a collection of documents.

Typical English language stop-words include “the”, “and”, “that”, “or”, “me”, etc. However, in order to improve the words considered by the algorithm, a list of custom stop-words was developed. Custom stop-words were evaluated by a subject matter expert in fire protection and iteratively introduced in the analysis. The analysis also used different algorithms to reduce the risk of bias affecting one algorithm. The analysis includes different algorithms and dimensionality reduction approaches which are discussed later in this section. The analysis also evaluates reports where a high fire protection loss occurred to ensure those reports were captured by the machine learning analysis. Since the analysis does not rely on structured data it can serve to complement conclusions obtained from the structured data. For example, the structured data can show that brush fires trends are increasing, the text analysis should show terms related to brush fires would also have an increasing trend.

This analysis uses the clustering algorithms Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and K-Means along with dimensionality reduction[7] approaches, i.e., (1) principal component analysis (PCA) and (2) t-distributed stochastic neighbor embedding (TSNE). At a high level, the clustering algorithms read the reports and group them together based on similar terms (i.e., assign them to a cluster). The DBSCAN algorithm requires the user to define several parameters. These include a distance/density value called Epsilon (EPS) and a minimum number of reports (defined as minimum samples in a cluster) to define a cluster. When reports are within the EPS distance, and meet the minimum samples criteria, they are grouped together. The K-Means algorithm requires the user to define the number of clusters (k). One of the main differences between K-Means and DBSCAN is that K-Means will assign a cluster to every single report while in DBSCAN assigns a report to a cluster if the algorithm criteria is met.

This section contains the results of the DBSCAN and K-Means clustering and text analyses. The reports were analyzed using a combination of DBSCAN, PCA and TSNE along with different approaches to vary the EPS value and minimum samples in a cluster to identify the terms. In the K-Means a k value of 10 was used. The clusters shown in the result figures in this section are those that resulted in an actionable term. These terms typically have a TF-IDF scoring of at least 0.1, are important to the cluster of reports and should be looked further by an ES&H subject matter expert and explored within the data. Even when the score is below 0.1, if the top 10 terms are related it can result in actionable terms. A term might also be important if it appears across multiple clusters.

The DBSCAN algorithm utilized the following three approaches to analyze the report text:

(1) Clusters with at least 6 reports using the Optimal EPS. The value of 6 reports represents the minimum samples in a cluster and approximately 1% of the reports.
(2) Cluster with the maximum number of reports when using the Optimal EPS. This cluster represented the most common terms within the reports collection.
(3) Clustering of 95% of reports. This approach iterates thru the EPS values until 95% of the reports are assigned a cluster. This approach also uses a minimum sample in a cluster of 6. The 5% of reports that were not assigned a cluster are outliers and assigned the label “Cluster 0”.

[7] Algorithms that convert multi-dimensional data into 2- or 3-dimensional space.

2.5.2 DBSCAN Clustering Algorithm with PCA

The Figure 30 thru Figure 34 in this section were the results of using the DBSCAN algorithm and PCA dimensionality reduction. Figure 31, Figure 33, and Figure 34 show the terms where the average score was at least 0.1. The following terms were identified as actionable and warranted further review (see Section 2.62.6): “capacitor”, cigarette related terms (“cigarette”, “receptacle”, “smoldering”, “receptacle smoldering”, “cigarette receptacle”, “smoldering outside”) and cold weather related terms (“sprinkler head”, “fan”, “froze cold”, “pipe”, “cold weather”) and “modulator”.

2.5.2.1 DBSCAN and PCA: Clusters with at least 6 reports and Optimal EPS

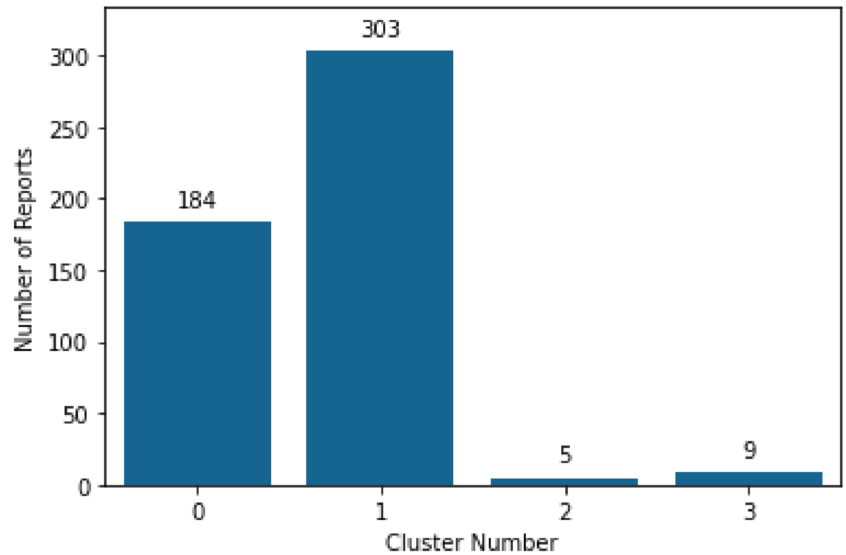

Figure 30: Number of Cluster for DBSCAN/PCA Approach 1

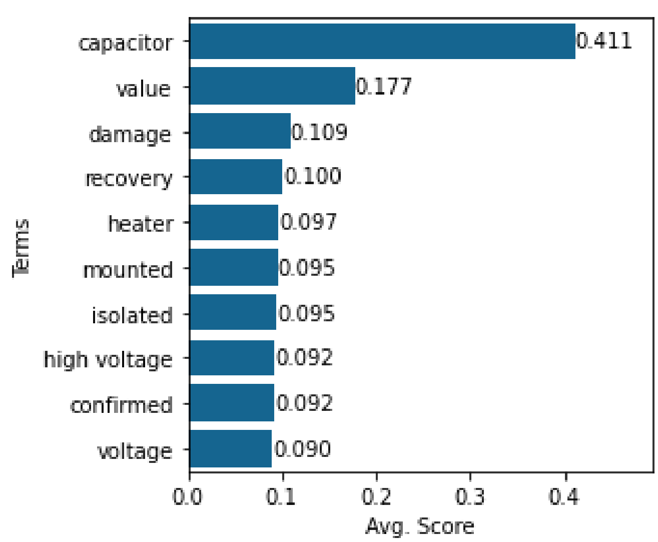

Figure 31: Term Avg. Score for Cluster 2

2.5.2.2 DBSCAN and PCA: Cluster with the maximum number of reports and Optimal EPS

Did not result in significant or new insights or terms.

2.5.2.3 DBSCAN and PCA: Clustering 95% of reports

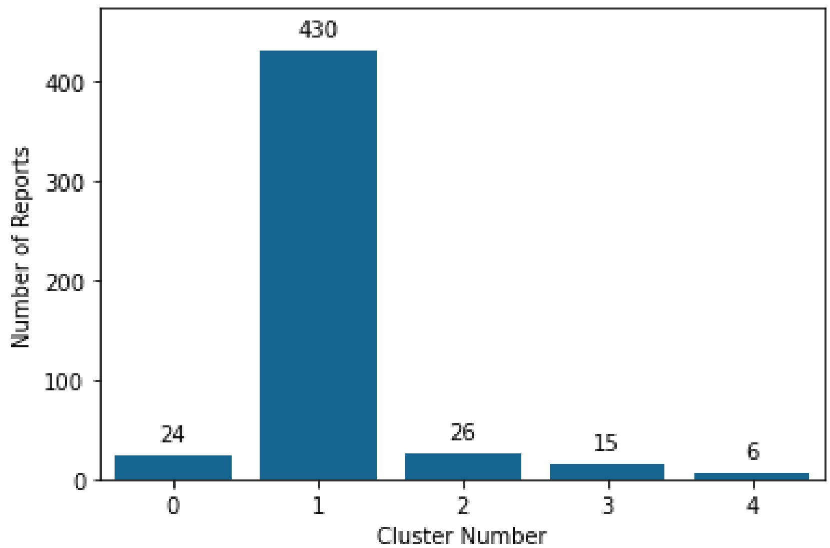

Figure 32: Number of Cluster for DBSCAN/PCA Approach 3

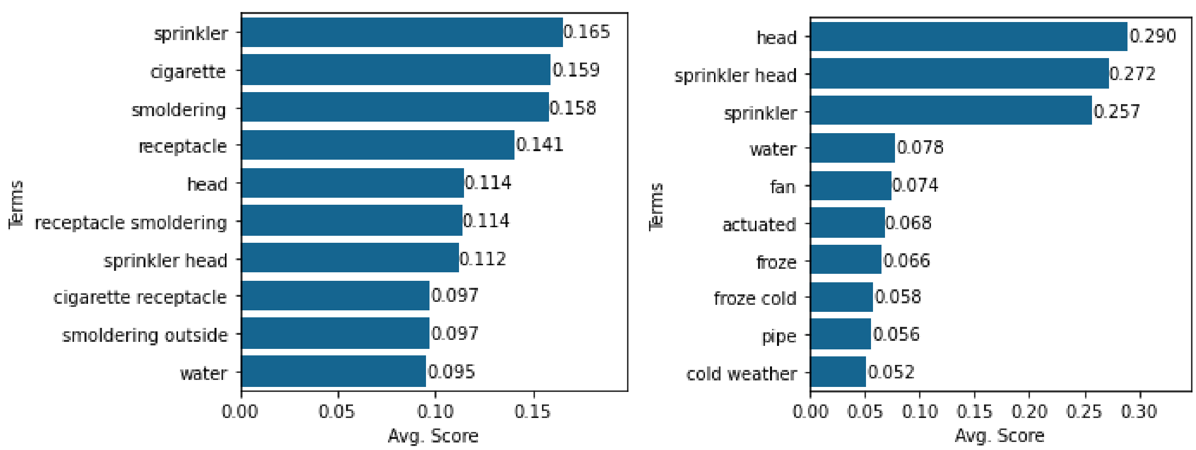

Figure 33: Term Avg. Score for Cluster 0 (left) and 2 (right)

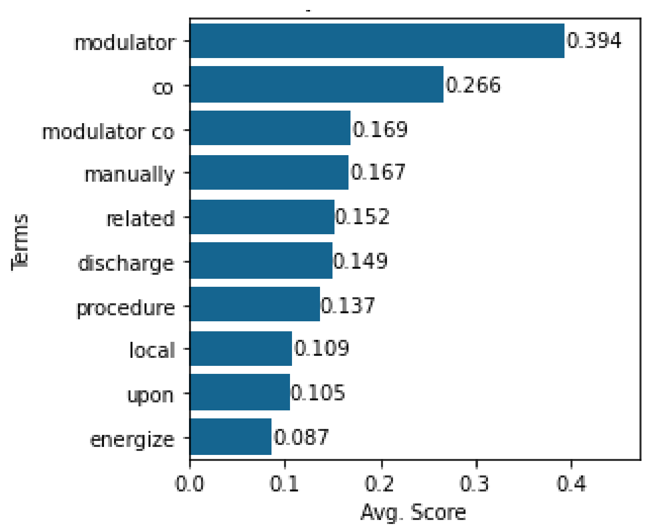

Figure 34: Term Avg. Score for Cluster 3

2.5.3 DBSCAN Clustering Algorithm with TSNE

The Figure 35 thru Figure 38Figure 30 in this section were the results of using the DBSCAN algorithm and TSNE dimensionality reduction Figure 36, Figure 37, and Figure 38 show the terms where the average score was at least 0.1 and not already identified in previous approaches. The following terms were identified as actionable and warranted further review (see Section 2.6): fume hood related terms ("fume hood", "hood", "fume", "chemical"), heating, ventilation, and air conditioning (HVAC) related terms ("belt", "hvac", "fan"), microwave related terms ("microwave", "left excessive", "burning", "smoking"), "transformer", and compressor related terms ("air compressor", "compressor").

2.5.3.1 DBSCAN and TSNE: Clusters with at least 6 reports and Optimal EPS

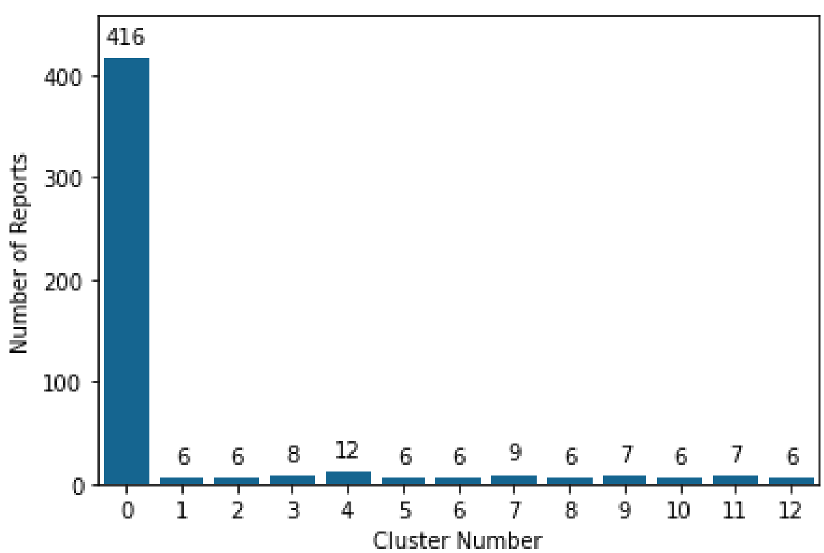

Figure 35: Number of Cluster for DBSCAN/TSNE Approach 1

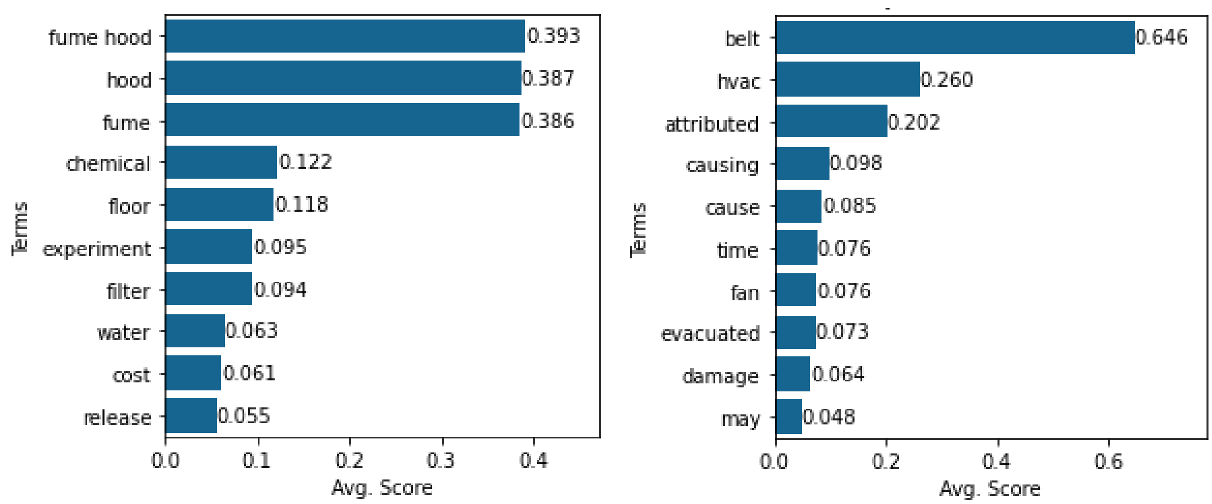

Figure 36: Term Avg. Score for Cluster 1 (left) and 2 (right)

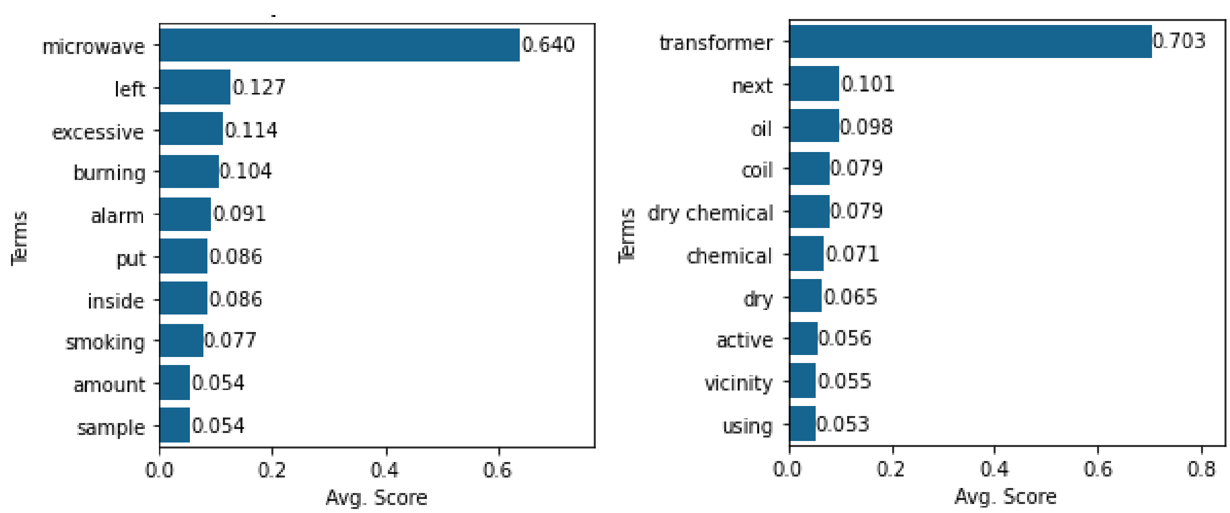

Figure 37: Term Avg. Score for Cluster 3 (left) and 5 (right)

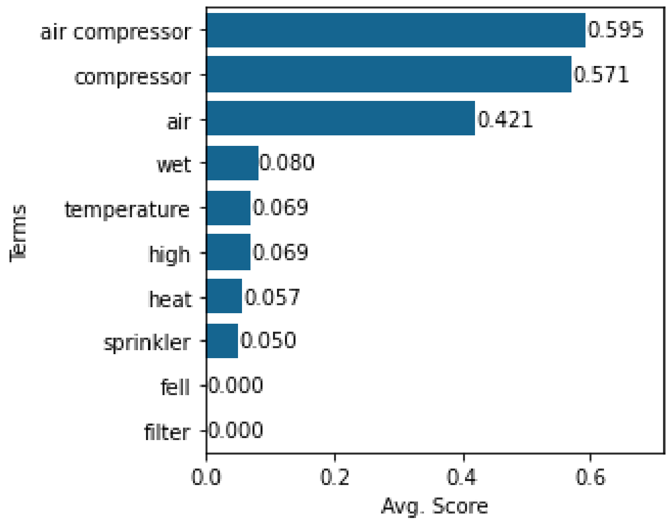

Figure 38: Term Avg. Score for Cluster 6

2.5.3.2 DBSCAN and TSNE: Cluster with the maximum number of reports and Optimal EPS

Did not result in significant or new insights or topics.

2.5.3.3 DBSCAN and TSNE: Clustering 95% of reports

Did not result in significant or new insights or topics.

2.5.4 K-Means Clustering Algorithm

Figure 39 and Figure 40 in this section were the results of using the K-means algorithm (with k = 10) and no dimensionality reduction. Figure 40 show the terms where the average score was at least 0.1 and not already identified in previous approaches. The following terms were identified as actionable and warranted further review (see Section 2.6): lightning strike related terms (“brush”, “lightning”, “lightning strike”, “strike”, “wildland”).

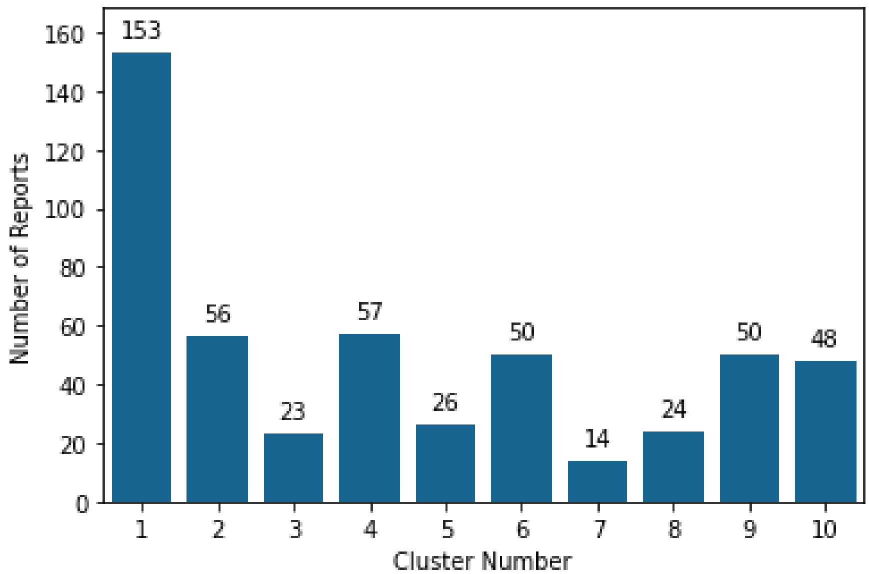

Figure 39: Number of Cluster for K-Means

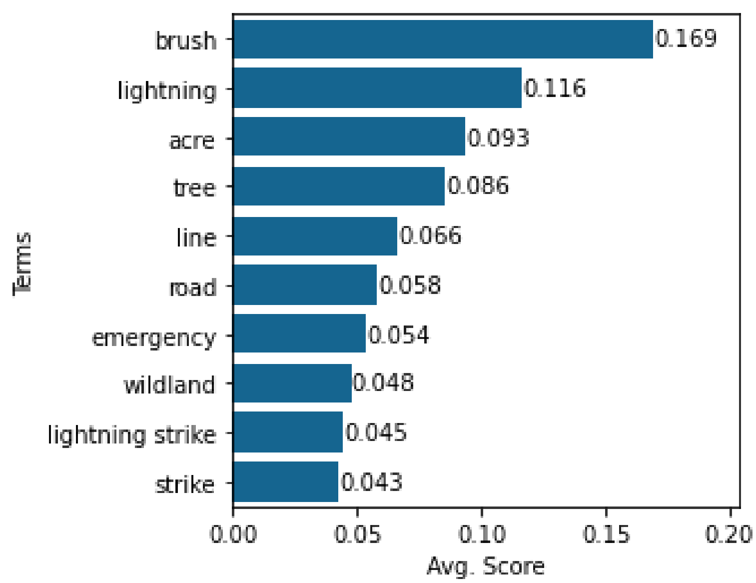

Figure 40: Term Avg. Score for Cluster 6

2.5.5 Text Analysis of Fire Protection Loss Reports Corpus

This section identifies the top 20 topics across the corpus of reports between 2015-2019. The analysis adjusted the TF-IDF vectorizer function maximum document frequency parameter (max_df)[8] to values of 0.95 and 0.05. A value of 1.0 does not ignore any vocabulary in the corpus. The value was adjusted iteratively (i.e., down to 0.05) until very common words that don't provide a lot of meaning in the fire protection loss data were ignored by the algorithm. Using this method, only the terms "fan", "pole" and "vehicle" were not identified by previous approaches using text clustering (e.g., "fan", "microwave", "lightning", "modulator").

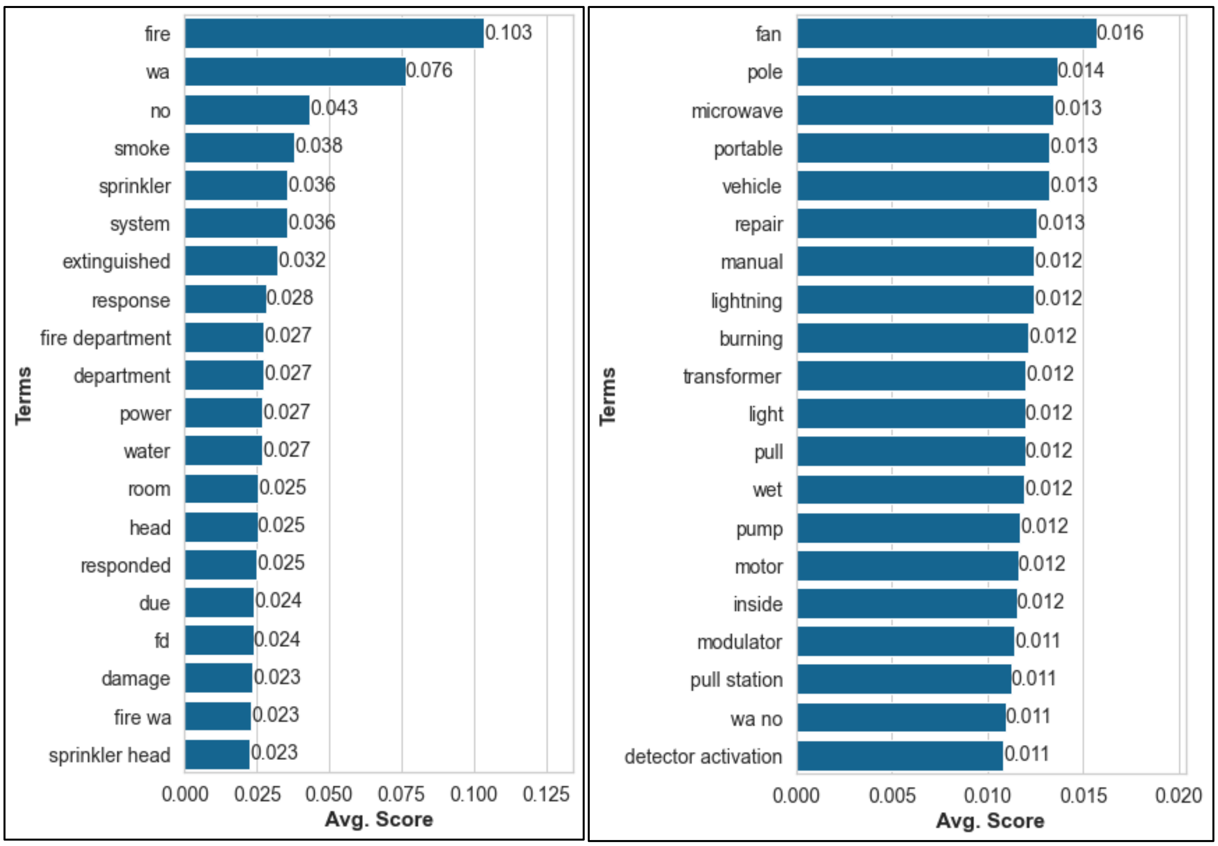

Figure 41: TF-IDF term average scores for max document frequency of 0.95 (left) and 0.05 (right)

[8] Standard practice in NLP and clustering is using values of max_df between 0.95 to 0.80.

2.6 Fire Protection Loss Reports Text Analysis

This section analyzes and discusses reports related to the terms identified in Sections 2.5.2 to 2.5.5. Table 4 provides a summary of the recurrent terms identified by all the text analysis and clustering methods and the average loss of the reports that resulted in significant losses. This table is sorted by the average losses of reports resulting in losses larger than $10K. The top terms resulting in large losses were related to expensive equipment (e.g., gas chromatograph, transformers, and HVAC), vegetation fires (e.g., from lightning-strike or downed power poles) and cold weather-related terms. These terms were used to filter and analyze the report data and obtain details and insights. Note that some terms might have overlapping fire protection reports. For example, the terms “pole” and “lightning” have reports associated to lightning strike induced fires that could have started by the lighting hitting a power pole or the resulting fire causing damage to power poles.

Table 4: Summary of recurrent terms and average losses

Terms	Number of related fire protection reports	Reports Resulting in Losses Larger than $10K		
		Number of Reports	Total Losses	Average Losses
Gas Chromatograph*	1	1	$750K	$750K
Lightning strike related terms (“brush”, “lightning”, “lightning strike”, “strike”, “wildland”)	23	2	$715K	$358K
“Transformer”	14	4	$894K	$223K
Cold weather-related terms (“sprinkler head”, “fan”, “froze cold”, “pipe”, “cold weather”)	20	3	$643K	$214K
“Pole”	20	5	$767K	$153K
Heating, ventilation, and air conditioning (HVAC) related terms (“belt”, “hvac”, “fan”)	16	4	$578K	$144K
“Capacitor”	15	2	$115K	$58K
Microwave related terms (“microwave”, “left excessive”, “burning”, “smoking”)	14	2	$94K	$47K
Fume hood related terms (“fume hood”, “hood”, “fume”, “chemical”)	10	3	$115K	$38K
“Fan”	20	3	$84K	$28K
“Vehicle”	19	7	$148K	$21K
Compressor related terms (“air compressor”, “compressor”)	10	3	$60K	$20K
“Modulator”	18	0	$0	$0
Cigarette related terms (“cigarette”, “receptacle”, “smoldering”, “receptacle smoldering”, “cigarette receptacle”, “smoldering outside”)	26	0	$0	$0
*This is a rare event identified by analyzing high monetary losses (Section 2.7).				

Fourteen fire protection loss reports were related to microwave related terms (“microwave”, “left excessive”, “burning”, “smoking”). Seven of these reports did not indicate monetary losses. Fire protection losses related to operator error (e.g., overheating, unknowingly leaving metallic parts, etc.) or microwave failure resulted in less than $1000 or minor injuries. A monetary loss of $57K was due to an analytical microwave due to the wrong probe being used and another loss of $37K in a microwave accelerated reaction system possibly due to equipment failure. The main cause of fire protection losses related to microwave seem to be operator error.

Fifteen fire protection loss reports were related to the term “capacitor” resulted. Eleven of these reports indicated monetary losses. All the “capacitor” related events are related to equipment failure. None of these events resulted in injuries potentially due to operators recognizing the risks of high voltage operations and clearing the area. One event resulted in more than $100K losses.

Eighteen fire protection loss reports were related to the term “modulator”. All of these reports indicated monetary losses of less than $10K. None of these events resulted in injuries. Causes were not determined but fire or damage was limited due to automatic or manually operated CO2 system being in place, in operation and activated. Fire protection losses seem to be due to actuation of the CO2 system and modulator damage.

Twenty fire protection loss reports were related to the term “pole”. None of the events indicated injuries. All events were related to equipment failure (e.g., electrical arcing), downed power lines, lightning strikes, which in many cases resulted in vegetation fires. Five events resulted in monetary losses of more than $10K. One event in 2019 caused by lightning resulted in INL’s largest fire with estimated loss of $700K. Another event in 2018 caused $200K losses due to a wildland fire from high voltage power supply lines.

Twenty-three fire protection loss reports were related to lightning strike related terms (“brush”, “lightning”, “lightning strike”, “strike”, “wildland”). Several of these 23 reports are related to losses caused by power poles or wildland fires potentially due to lighting strikes. The INL fire (discussed previously) had an estimated monetary loss of $700K, another report had a loss of $15K, all other losses were below $5K.

Fourteen fire protection loss reports were related to the term “transformer”. These fires were mostly limited to the transformer unit with the main cause being electrical failure. Five events had a monetary loss greater than $5K. These included monetary losses of $434K (CY2015), $280K (CY2016), $80K (CY2017), $72K(CY2015), $26K (CY2017).

Twenty fire protection loss reports were related to cold weather-related terms (“sprinkler head”, “fan”, “froze cold”, “pipe”, “cold weather”). None of these events resulted in injuries. Two events in 2015 resulted in $500K (ORPS reportable), $133K of losses and $10K. Both events were due to piping freezing and breaking resulting in building flooding. All other events resulted in less than $10K losses. Main causes reported were improper building heating and equipment failure due to cold weather conditions.

Twenty fire protection loss reports were related to the term “fan”. Seventeen of the 20 reports resulted in monetary losses of less than $10K. None resulted in injuries. The most common issue was activation of sprinkler systems due to high temperature caused by a fan failure or a fire resulting from fan equipment failure.

Ten fire protection loss reports were related to compressor related terms (“air compressor”, “compressor”). Eight of the 10 reports resulted in monetary losses of $30K or less. All events except one were caused by equipment failure. One of these reports discusses a maintenance issue were a dirty filter of an additive manufacturing machine compressor/recirculation pump caught fire because the filter was not maintained because it was difficult to access and not described in the operating manual. This event resulted in $7K loss. None of these events resulted in injuries.

Ten fire protection loss reports were related to fume hood related terms (“fume hood”, “hood”, “fume”, “chemical”). Seven of the 10 reports resulted in monetary losses up to $80K. All events were caused by chemical reactions. One event resulted in minor injuries when vessels containing the chemicals injured the employee. This event is discussed in more detail in Section 2.7.

Twenty-six fire protection loss reports were related to cigarette related terms (“cigarette”, “receptacle”, “smoldering”, “receptacle smoldering”, “cigarette receptacle”, “smoldering outside”). One of these reports resulted in a $2500 loss while the rest were below $250. All were small fires related to vegetation, trash can or cigarette receptacles.

Nineteen fire protection loss reports were related to the term “vehicle” . Seventeen of reports were related to vehicle fires. When reported the most common cause was electrical failure.

Sixteen fire protection loss reports were related to heating, ventilation, and air conditioning (HVAC) terms (“belt”, “hvac”, “fan”). Four of these reports were related to HVAC fires and resulted in monetary losses of $398K (CY2015), $105K (CY2019), $60K (CY2019) and $15K (CY2016). All other HVAC related fire protection losses were less than $5K. The largest fire protection loss HVAC-related fire occurred in 2015 and the possible cause was determined to be incorrect installation of an HVAC unit by the trailer manufacturer and degradation of the trailer wood frame. The $60K fire occurred in the same trailer, but in 2019. The origin of the fire was determined to be in the vicinity of the HVAC unit. After investigation the cause was unspecified. The only recurring cause reported was equipment failure.

The text analysis in this section could be expanded to include trends of the topics over the period as well as distribution of incidents per site. In a similar way to the trends of the structured data, this can provide insights on topics that have increasing or decreasing trends and can be used to complement observations of the structured data trends. The text analysis can also be expanded to other report sets such as the fire protection ORPS reports discussed in Section 3. This could provide some insights into recurrent terms that meet the ORPS threshold and compared to the terms of the fire protection loss reports discussed in this section. However, this would significantly expand the report and the intent of the section was to provide high level insights on recurrent terms of the fire protection database.

2.7 Analysis of Fire Protection Reports with Hight Monetary Losses or Injuries

This section discusses individual events that resulted in high monetary loss (i.e., larger than $10K) and are not discussed in Section 2.6. One explanation on why these reports were not identified by the machine learning algorithm is because the number of related reports did not meet the minimum samples in a cluster (e.g., 6). By definition these are low frequency incidents, nonetheless, they resulted in large monetary losses or injuries and are evaluated below. Table 5 provides a summary of the terms and the average loss of these reports.

Table 5: Summary of terms not captured in the text and cluster analysis

Terms	Number of related fire protection reports	Number of reports resulting in significant losses	Average loss of reports resulting in significant losses
Gas Chromatograph	1	1	$750K
Injuries	5	Injuries are not quantified	

The largest loss over the 2015-2019 calendar year period was related to a 2019 gas chromatograph fire where its associated thermal desorption unit failure resulted in $750K damage. It is important to note that there were no other reports in the data involving a chromatograph. The chromatograph report is a very low frequency (i.e., rare event) with high consequence. Rare events are difficult to detect using short time frames (e.g., 5 years in this case). These machine learning text analysis techniques have limitations detecting rare events as they rely on frequency of terms. A risk analysis in specific areas could provide insights into risk of some operations and complement trending and machine learning analysis methods.

Other related causes of high monetary losses ($200K or less) included equipment failures and material related fires. Five reports discuss injuries. One of five reports was related to minor burns to the hand/fingers, and two of the reports discuss injuries caused to exothermic chemical reactions in a fume hood and the debris from the vessel containing the chemicals injuring a person.[9] One of the injury related reports was mis-classified. However, the text and clustering analysis helped identify this error. Another report had an injury to an emergency response personnel classified as unspecified.

2.8 2011-2015 Fire Protection Trend Analysis and Verification of the Text Analysis

To ensure that the results of the machine learning analysis were trustworthy and were unlikely to miss any important topics the authors reviewed the findings of a previous fire protection trending analysis of 2011 to 2015 fire protection data. That analysis recommended new fire protection loss cause categories and topics that were frequently occurring. These recommended new causes categories include:

- Human Carelessness/Vandalism
    - Discarded cigarette/cigar butts
    - Kitchen/microwave fires
    - Plug-in heaters
    - Daisy-chained electric cords/power strips
    - Fireworks/smoke bombs
- Maintenance Issues
    - Deferred maintenance
    - Missed preventive maintenance
    - Missed surveillance
    - Clogged vents

[9] ORPS Report NA--LASO-LANL-RADIOCHEM-2015-0001. Caused by an inadvertent container exothermic reaction during closing-down operations. Employee was wearing safety glasses and minor injuries included lacerations in hand and forehead. ORPS Report NA--LASO-LANL-NUCSAFGRDS-2018-0002. Caused by a ruptured vessel containing a chemical reaction. The employee received minor injuries in the hands, arms, neck, and face.

    - Overfilled cigarette butt cans that catch fire and emit smoke
- Animal damage
    - Chewed wires
    - Nesting in vents.
    - Electrical wire contact
    - Defecation damage

To verify the results of the text analysis, the authors performed word searches on the description of the 2015-2019 fire protection loss reports. Two topics that seemed recurrent in the previous analysis and in this one, are those terms related to microwaves and cigarettes. As stated below, the other terms did not result in any frequent, relevant, or high-cost reports confirming that the text and machine learning analysis did not miss any important issues identified in the previous trending period.

No fire protection loss reports were found to be related to the terms “firework”, “bomb”, “daisy”, “chained”, “clog”, and “chew”. Two reports were related to the term “heater”, specifically to a “plug-in space heater”, none resulted in significant losses.

Five fire protection loss reports were related to the term “maintenance”. None of the reports specified that the cause of the loss was due to deferred or missed maintenance.

Tow fire protection loss reports were related to the term “surveillance”, none related to “missed surveillances”.

Two fire protection loss reports were related to the term “nest”. Only one report was related to a fire from a rodent nest that resulted in $5K damage.

Two fire protection loss reports were related to the term “animal”. Both reports determined that it was possible that an animal causing a short resulted in igniting vegetation and both resulted in zero losses. Other animal related words “defecation”, “bird”, “bear”, “deer”, resulted in no reports. Animal related fire protection losses don’t seem to be frequent during this period.

3 ORPS Fire Protection Reports

The ORPS database captures safety related occurrences that meets its program criteria. Fire protection related reports are assigned ORPS HQ keyword 03 "Fire Protection and Explosives Safety." This section discusses the trends of ORPS fire protection related reports. Note that some reports could meet the reporting criteria of both the ORPS and the FPDB while other reports only meet the reporting criteria of one and are captured within one database only.

This section discusses trends and distribution of keywords of the ORPS fire protection reports. The Figure 42 shows the trends for ORPS fire protection related reports. It shows a decreasing trend between 2015 and 2019. During this timeframe there were a total of 376 ORPS fire protection reports.

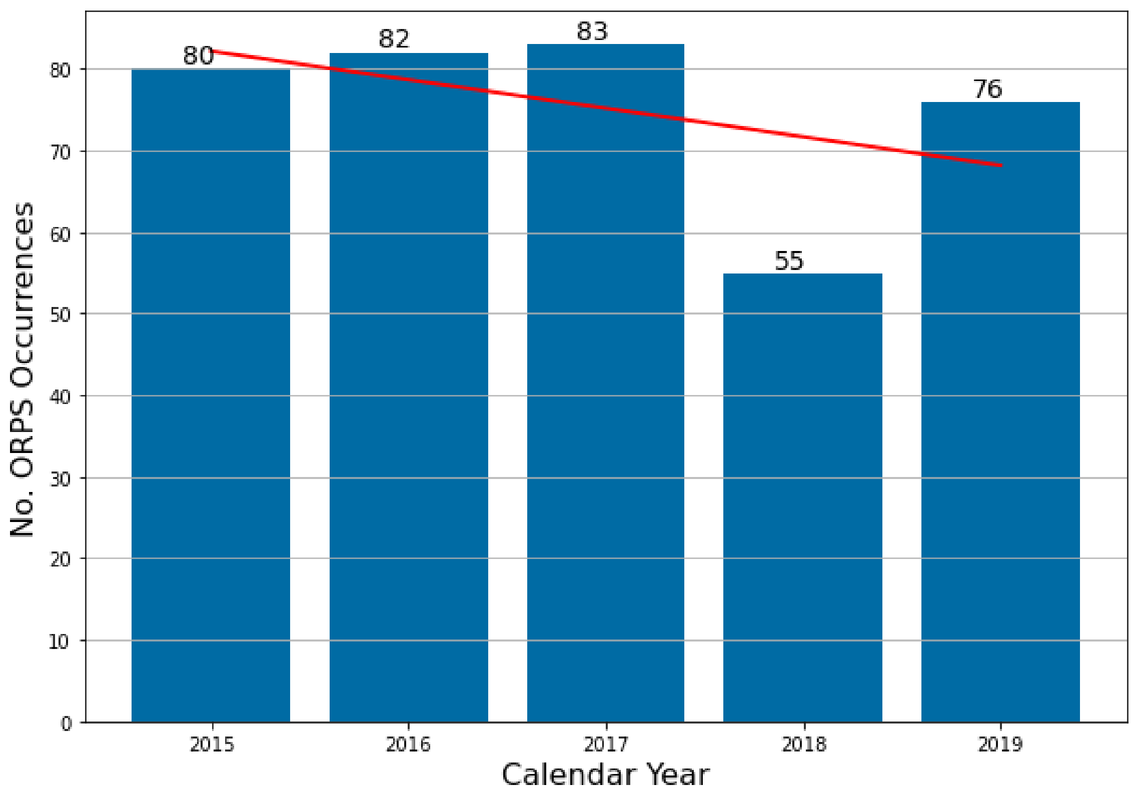

Figure 42: ORPS Fire Protection Related Reports (Keyword category 03) per Calendar Year

Table 6 shows the top 15 keyword distribution of reports (excluding categories 12-EH Categories and 14-Quality Assurance). This table shows that the most common reports were those related to 01-Work Planning and Control Deficiencies, 01R-Management Issues with 60 reports, 01B-Loss of Configuration management/Control with 51 reports and 01E-Facility Operations Procedure Noncompliance with 38 reports being the top keywords related to causes. Note that keyword related to Inadequate Conduct of Ops was retired during the 2017 ORPS DOE Order revision. Of the non 01 related keywords, 05D-Mechanical Equipment Failure/Damage was the most common with 49 reports. There was a total of 216 reports related to fire protection equipment degradation.

Note that the revisions to DOE Order 232.2A, *Occurrence Reporting and Processing of Operations Information,* that came into effect in calendar year 2017, could have caused in part the decrease observed in the ORPS reportable occurrences in 2018 and 2019.

Table 6: Distribution of the top 15 ORPS keywords

ORPS keyword	Number of Reports
03A- Fire Protection Equipment Degradation	216
03C-Facility Fire	109
01I-Safety System Actuation/Evacuation	86
01R-Management Issues	60
01B-Loss of Configuration management/Control	51
05D-Mechanical Equipment Failure/Damage	49
01A-Inadequate Conduct of Ops (Retired)	49
01E-Facility Operations Procedure Noncompliance	38
01N-Inadequate Job Planning (Other)	35
01G-Inadequate Procedure	31
11G-Subcontractor	31
07E-Electrical Equipment Failure	29
01P-Inadequate Oral Communication	27
03D-Explosive Safety Issue	27
01Q-Personnel Error	25

Table 7 shows the top 15 combined keyword distribution. This table shows the 03-Fire Protection keyword combinations. Of the 109 reports related to a facility fire (i.e., 03C), three are related to 03A-Fire Protection Equipment Degradation. There were 19 ORPS reports related to 03G-Wildland Fires.

Table 7: Distribution of the top 15 ORPS 03-Fire Protection keyword combinations

ORPS 03-Fire Protection Keywords Combination	Number of Reports
03A-Fire Protection Equipment Degradation	211
03C-Facility Fire	94
03D-Explosive Safety Issue	26
03G-Wildland Fire	18
03B-Fire Suppression Actuation and 03C-Facility Fire	9
03B-Fire Suppression Actuation	4
03F-Explosion	3
03E-NFPA Code Fire Protection Issue	3
03A-Fire Protection Equipment Degradation and 03C-Facility Fire	2
03A-Fire Protection Equipment Degradation and 03B-Fire Suppression Actuation	2
03A-Fire Protection Equipment Degradation and 03B-Fire Suppression Actuation and 03C-Facility Fire	1
03C-Facility Fire and 03G-Wildland Fire	1
03C-Facility Fire and 03F-Explosion	1
03C-Facility Fire and 03D-Explosives Safety Issue	1

The ORPS reporting level (e.g., High, Low, and Informational) provide a means to reflect perceived risk associated with a given report. They take into consideration the potential consequence of an occurrence in terms of health, safety and security to personnel, the public, the environment, and the operational mission. Of the 376 total reports between 2015 and 2019 in the ORPS "03-Fire Protection" report data, 57 were High, 141 were Low, and 178 were at the Informational reporting level.

Table 8 shows the number of reports related to 14-Quality Assurance. In this case a report has to have at least one assigned but can have multiple keywords assigned. Of the 376 reports, 218 found no QA deficiency (i.e., 14L). Keyword 14E-Work Process Deficiency had the greatest number of reports with 136 followed by "14D-Documents and Records Deficiency" with 70 reports.

Table 8: Distribution of ORPS Fire Related Reports Quality Assurance Keyords

ORPS QA keyword	Number of Reports
14L-No QA Deficiency	218
14E-Work Process Deficiency	136
14D-Documents & Records Deficiency	70
14G-Procurement Deficiency	27
14B-Training & Qualification Deficiency	22
14C-Quality Improvement Deficiency	17
14H-Inspection & Acceptable Testing Deficiency	16
14F-Design Deficiency	14
14A-Program Deficiency	5

4 Fire Department Response Data

This section discusses trending data from fire department response calls. The DOE wide, calendar year total fire department response calls can be found in Table 9 and includes calls related to fires, hazardous materials (HazMat), medical, non-emergency, and other emergencies. The data in Table 10: DOE Wide Total Number of Fire Department Response Calls and Normalized Values, shows the total number of fire department response calls per CY, the total work hours (for sites that report into fire department response call data) and the normalized total number of fire department response calls.

Table 9: DOE Wide Fire Department Response Data

Calendar Year	Number of Fire Calls	Number of HazMat Calls	Number of Medical Calls	Number of Non-Emergency Calls	Number of Other Emergency Calls	Total Number of Calls
2015	357	225	1,624	2,333	1,701	6,240
2016	404	257	1,665	2,161	1,637	6,124
2017	476	222	1,569	2,104	1,477	5,844
2018	412	203	1,924	2,277	1,673	6,489
2019	460	265	2,042	2,412	1,785	6,964

Table 10: DOE Wide Total Number of Fire Department Response Calls and Normalized Values

Year	Total Number of Calls	Total Work Hours	Normalized Total Number of Calls
2015	6,240	193,617,886	6.45
2016	6,124	199,289,648	6.15
2017	5,844	203,222,276	5.75
2018	6,489	204,418,582	6.35
2019	6,964	219,683,923	6.34

Figure 43 shows the number of fire department response calls trend per CY. Even though this trend is slightly increasing at a pace of approximately 181 calls per year, the normalized total number of fire department response calls (Figure 44) shows a flat trend which means that the total work hours are also increasing at a proportional pace. Section 3, above, shows that the ORPS reports related to fire protection (Figure 42) have been decreasing. However, no correlation between ORPS and fire department response data was determined.

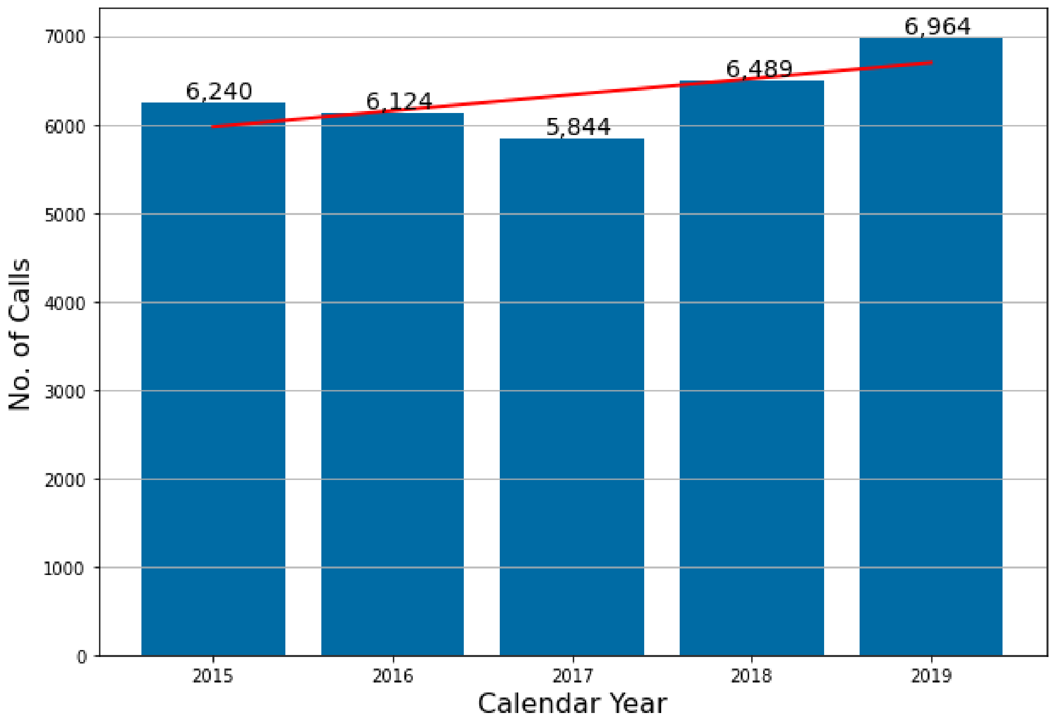

Figure 43: Total Number of Fire Department Response Calls

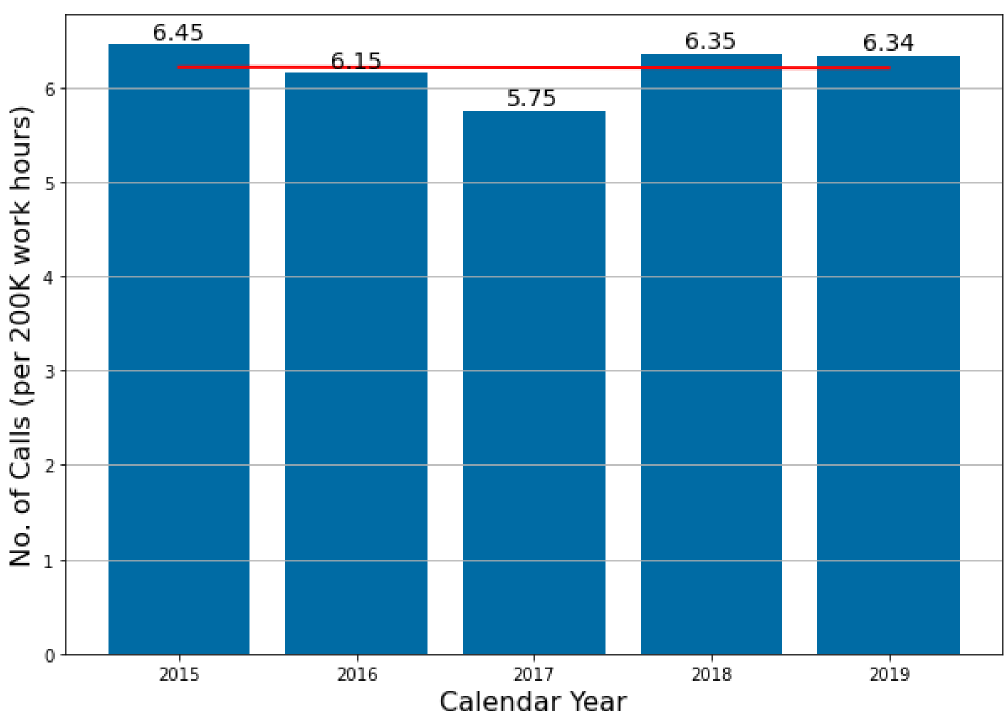

Figure 44: Total Number Fire Department Response Calls (per 200K work hours)

4.1 Site Fire Response Data

Several criteria can be used to compare response call trends at the site level, including normalized and non-normalized fire department responses and trends over a specific time period. The tables in this section were developed from an automated calculation where the slope of each site's trend was extracted in order to quickly identify the largest increasing trends. The same approach can be used with other types of calls (normalized and non-normalized) to calculate and evaluate the highest increasing trends, and sites can request additional tables from the Office of ES&H Reporting and Analysis.

Table 11 includes the 2019 calendar year total number of fire department calls sorted in descending order by the normalized total number of calls. This table shows that in 2019, sites with an internal fire department tended to report more fire department response calls per 200K work hours than sites that use a public (e.g., city or county) fire department and emergency services. The data for other years shows the same characteristics but were not included in this report. This shows that there could be reporting differences for sites that rely in internal emergency services versus those sites that rely on public services. Although there are no firm explanations as to why this difference occurs it should be considered when site comparisons are being performed using fire department response data. In addition, the number and types of fire department and emergency calls that are reported at a site will vary based on the nature of operations and the work hours at that site. The section below uses representative sites to illustrate trends and to highlight some limitations and challenges of the data which include:

- Missing data for some sites
- Differences in reporting from year to year
- Differences in reporting from site to site

These observations were not specific only to the selected sites as similar observations were found across most of DOE sites. Addressing these observations could potentially result in different trends. AU is working to improve data collection, data quality, ability to measure trends and identify areas of improvements and potential leading indicators that could enhance learning and safety across DOE. Improvements will be coordinated with the users of the DOE's FPDB as part of a separate initiative and are not part of the scope of this report.

Table 11: 2019 Calendar Year Number of Calls data

Site	Number of Calls	Work Hours	Normalized Number of Calls (calls/200K work hours)	Fire Department services[10]
Princeton Plasma Physics Laboratory	396	961,262	82.39	Site
Paducah Gaseous Diffusion Plant	62	804,451	15.41	Site
Argonne National Laboratory	488	6,794,717	14.36	Site
Brookhaven National Laboratory	357	4,992,292	14.30	Site
Pantex Site	488	6,995,031	13.95	Site
Oak Ridge National Laboratory	629	9,557,931	13.16	Public on Site
Fermi National Accelerator Laboratory	261	4,090,987	12.76	Site
Portsmouth Gaseous Diffusion Plant	256	4,355,562	11.76	Site
Waste Isolation Pilot Plant	119	2,248,207	10.59	Site
Nevada National Security Site	198	5,000,540	7.92	Site
National Renewable Energy Laboratory	181	4,750,860	7.62	Public
Richland Operations Office	791	21,248,395	7.45	Site
Savannah River Site	721	20,579,163	7.01	Site
Y-12 National Security Complex	463	13,853,811	6.68	Site
Lawrence Livermore National Laboratory	440	14,201,274	6.20	Public on Site
Los Alamos National Laboratory	571	21,035,414	5.43	Public
Strategic Petroleum Reserves	36	1,959,923	3.67	Combination
Idaho National Laboratory	181	12,549,927	2.88	Site
East Tennessee Technology Park	50	3,899,316	2.56	Public on Site
West Valley Demonstration Project	8	634,500	2.52	Public
Lawrence Berkeley National Laboratory	88	7,220,836	2.44	Public on Site
Thomas Jefferson National Accelerator Facility	9	1,287,614	1.40	Public
SLAC National Accelerator Laboratory	22	3,150,725	1.40	Public
Office of Secure Transportation	7	1,510,863	0.93	AFB
Kansas City National Security Campus	36	8,573,270	0.84	Public
Sandia National Laboratories-New Mexico	95	26,132,423	0.73	Site
Ames Laboratory	2	713,974	0.56	Public
Pacific Northwest National Laboratory	9	8,095,801	0.22	Public
Office of Legacy Management	0	1,245,927	0	Indetermined
Oak Ridge Institute for Science and Education- Oak Ridge Associated Universities	0	1,238,927	0	Public

[10] Fire Department and Emergency Services provided by a site station, public (i.e., city or county station), air force base (AFB)).

Table 12 shows the trends[11] of the total number of fire department calls between 2015 and 2019 sorted by the trend of the normalized number of calls.

Table 12: Trends of the Total Number of Fire Department Responses

Site	Normalized Number of Calls Trend (calls/200K work hours) per year)	Number of Calls Trend Slope (calls/year)	Fire Department services
Princeton Plasma Physics Laboratory	6.76	25.3	Site
Paducah Gaseous Diffusion Plant	2.28	6.1	Site
National Renewable Energy Laboratory	1.73	41.5	Public
Portsmouth Gaseous Diffusion Plant	1.1	22.25	Site
Savannah River Site	0.82	91.7	Site
Waste Isolation Pilot Plant	0.49	8.9	Site
Pantex Site	0.46	27.7	Site
Los Alamos National Laboratory	0.44	54.1	Public
Brookhaven National Laboratory	0.27	-2.5	Site
West Valley Demonstration Project	0.24	0.9	Public
Y-12 National Security Complex	0.14	17.2	Site
Thomas Jefferson National Accelerator Facility	0.13	0.8	Public
Pacific Northwest National Laboratory	0.06	2.3	Public
Sandia National Laboratories-New Mexico	0.01	4.1	Site
Lawrence Livermore National Laboratory	0.01	18.6	Public on Site
Ames Laboratory	-0.03	-0.2	Public
Oak Ridge Institute for Science and Education-Oak Ridge Associated Universities	-0.03	-0.3	Public
Office of Legacy Management	-0.04	-0.2	Indetermined
Idaho National Laboratory	-0.05	6.8	Site
SLAC National Accelerator Laboratory	-0.06	-0.3	Public
Fermi National Accelerator Laboratory	-0.12	1.7	Site
Nevada National Security Site	-0.17	3.9	Site
East Tennessee Technology Park	-0.2	1.2	Public on Site
Kansas City National Security Campus	-0.22	-2.4	Public
Richland Operations Office	-0.64	-49.6	Site
Strategic Petroleum Reserves	-0.64	-4.5	Combination
Lawrence Berkeley National Laboratory	-0.72	-22.6	Public on Site
Argonne National Laboratory	-1.04	-27.6	Site
Oak Ridge National Laboratory	-1.17	-44.9	Public on Site
Office of Secure Transportation	NaN	NaN	AFB

[11] Trends: the trend number represent on average how much the number is changing (increasing/decreasing) per year. It represents the slope of the linear trendline.

4.2 Fire Response Data Trend Analysis at Selected Sites

The following representative sites were selected using the results shown in Table 11 and Table 12: Princeton Plasma Physics Laboratory (PPPL), Paducah Gaseous Diffusion Plant, and Argonne National Laboratory (ANL). These are the sites with the top three highest 2019 normalized number of calls (Table 11). Coincidentally, PPPL and Paducah have the two highest increasing trends for the normalized total number of calls with the National Renewable Energy Laboratory (NREL) having the third highest increasing trend. ANL had one of the lowest decreasing normalized total number of calls trend.

4.2.1 Princeton Plasma Physics Laboratory (PPPL)

Table 12 showed that PPPL had the highest increasing trend for the normalized total number of calls. Figure 45 show the trends for the total number of calls per CY and Figure 46 the normalized values. Both figures show that calls received have been constantly trending upward. In 2014, PPPL made significant changes that improved the consistency of reporting the number of their fire department response calls.

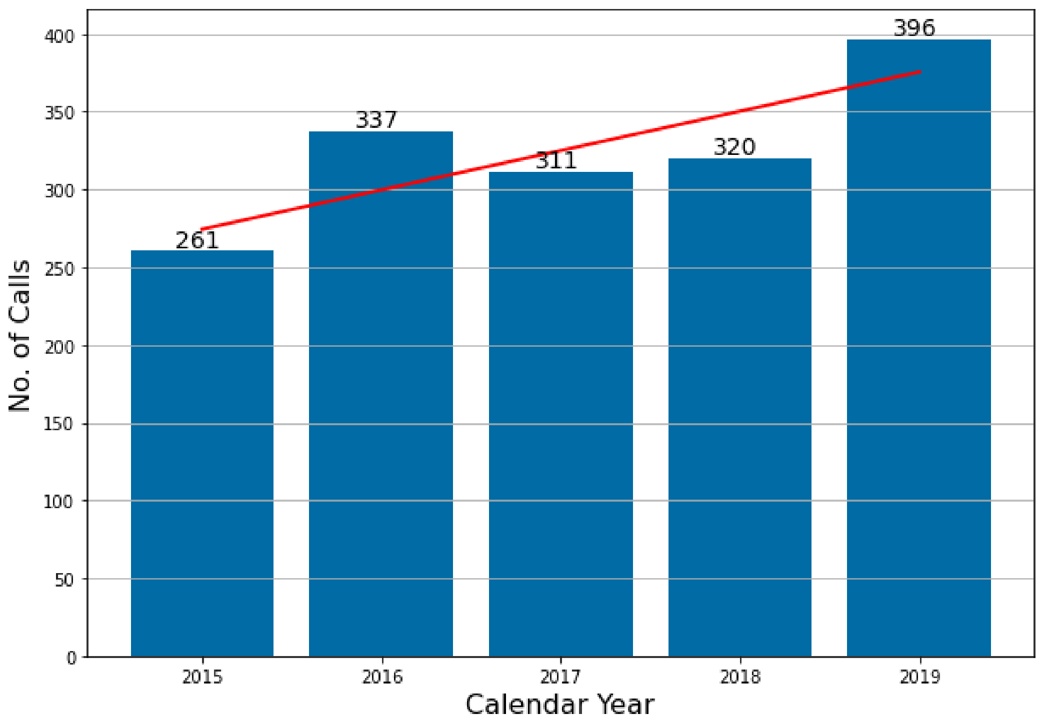

Figure 45: PPPL Number of Fire Department Response Calls

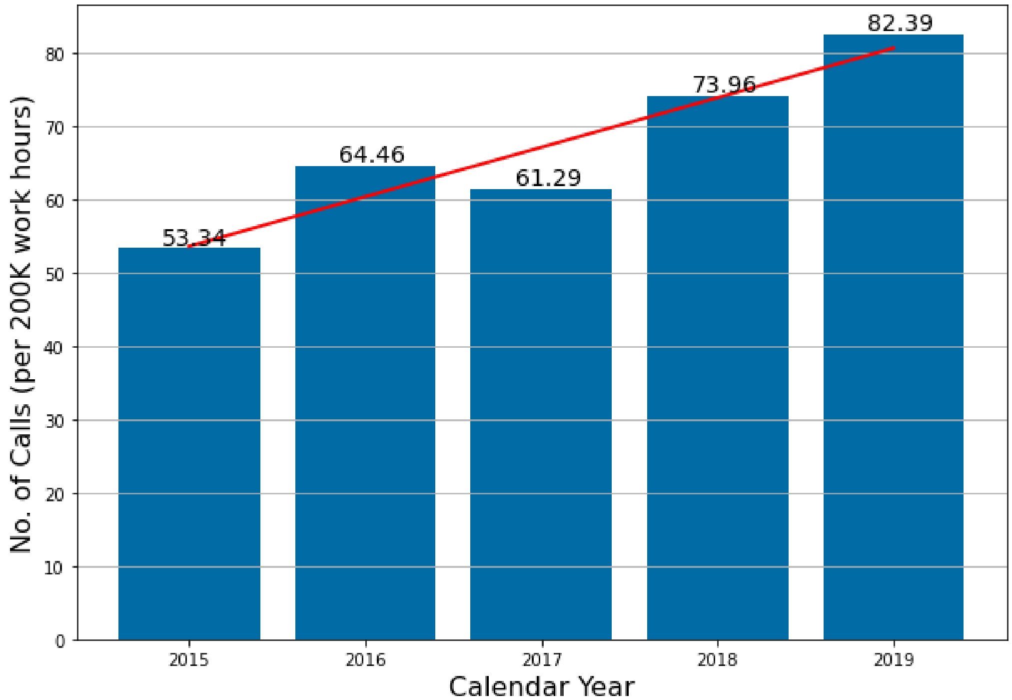

Figure 46: PPPL Number of Fire Department Response Calls (per 200K site work hours)

Figure 47 shows the distribution of the types of calls. Even though some of the trends are increasing (such as the number of other emergency calls), this figure shows that PPPL reporting is relatively consistent year over year and does not seem to have spikes in the data or missing data points.

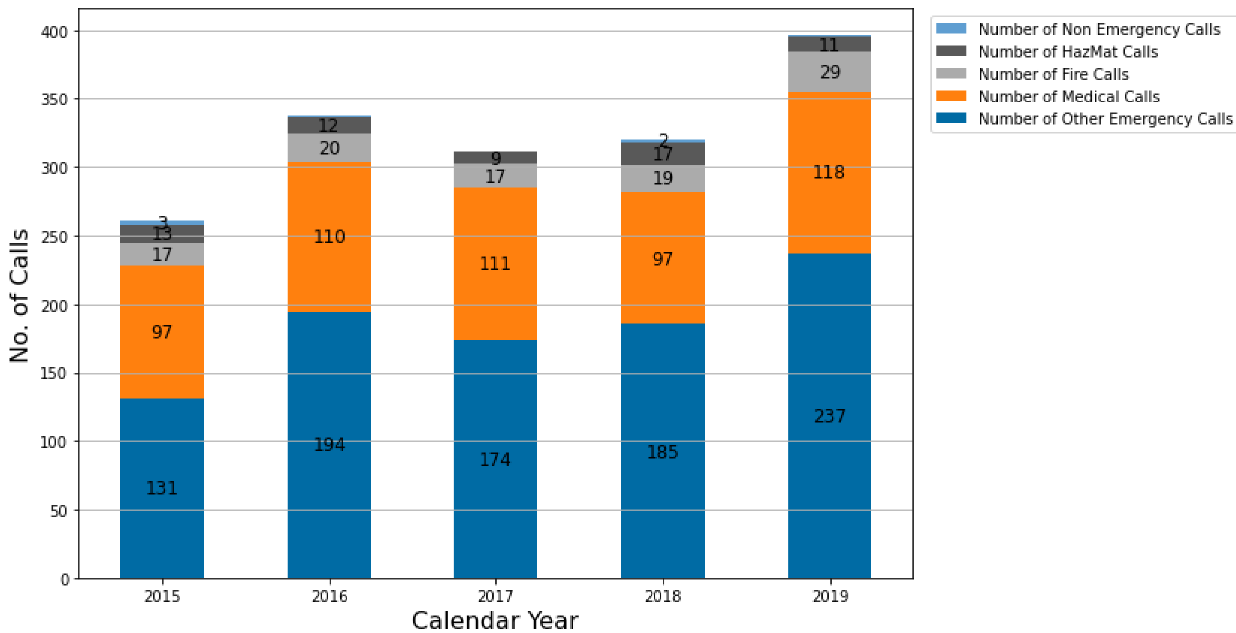

Figure 47: PPPL Fire Department Response Calls by Type

4.2.2 Paducah Site

Table 12 showed that Paducah site had the second highest increasing trend for the normalized total number of calls. Figure 48 show the trends for the total number of calls per CY and Figure 49 shows the normalized values.

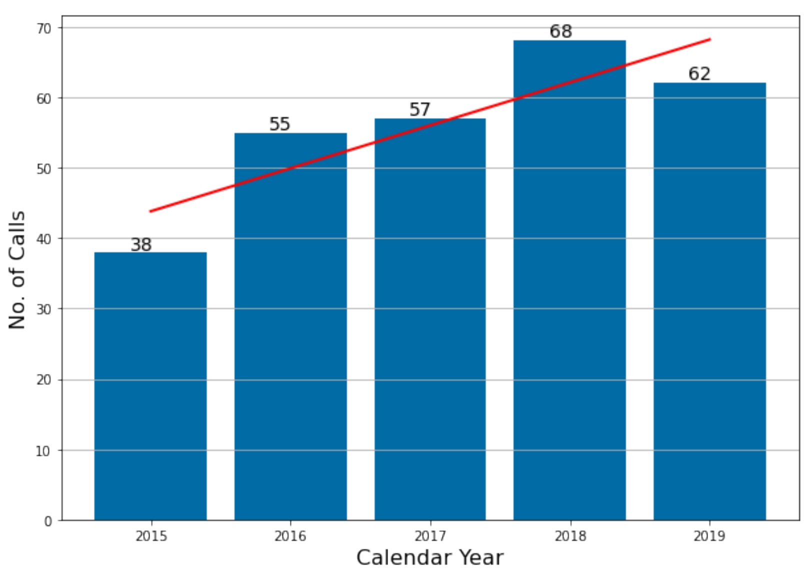

Figure 48: Paducah Site Number of Fire Department Response Calls

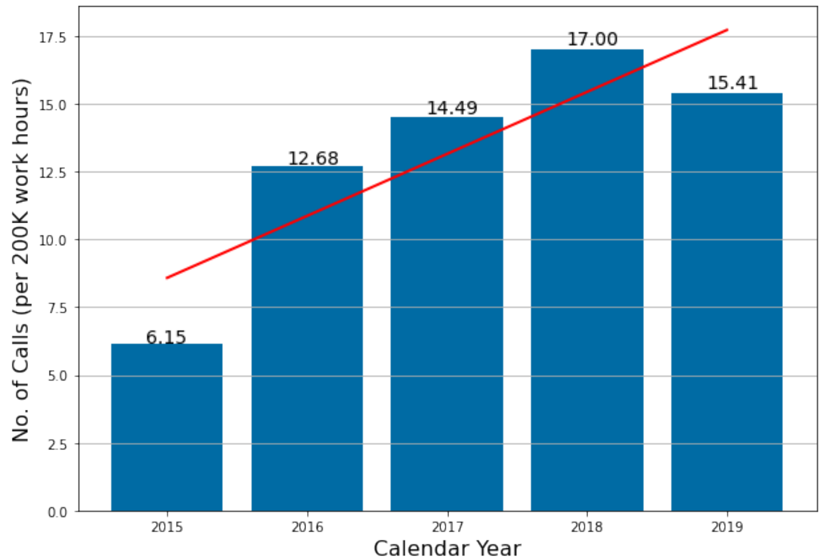

Figure 49: Paducah Site Number of Fire Department Response Calls (per 200K site work hours)

Figure 50 shows the distribution of the types of calls. The number of other emergency calls in the dataset varied from one call in 2015, zero calls in 2016, 2017, 2018 and 16 calls in 2019. This caused the trend for the other emergency calls to be the highest, followed by non-emergency and hazmat calls both which also show increasing trends. The trend for fire calls is flat and the medical calls are decreasing.

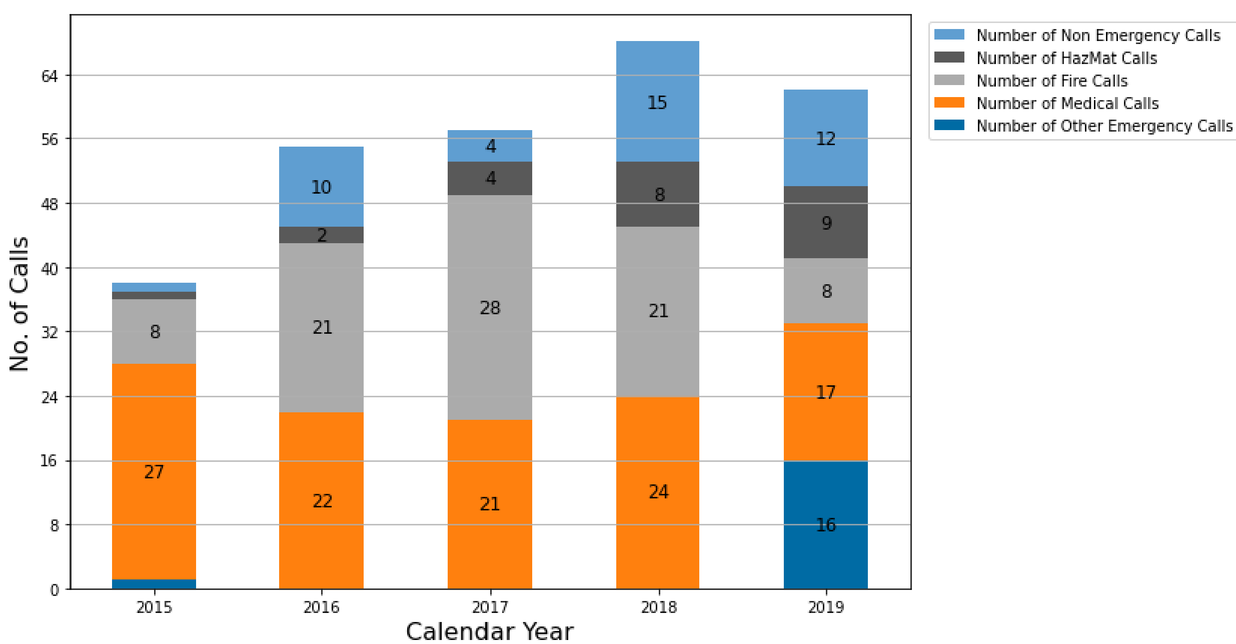

Figure 50: Paducah Site Fire Department Response Calls by Type

4.2.3 National Renewable Energy Laboratory (NREL)

Table 12 showed that the NREL site had the third highest increasing trend for the normalized total number of calls. Different from the other selected sites in the report, NREL uses a public fire department. Table 11 also shows that the 2019 normalized total calls value is much lower than many other sites. Figure 51 shows the trends for the total number of calls per CY and Figure 52 the normalized values. Both of these figures show that the trends are increasing, and both show a spike in the 2018 reported values. When analyzing the types of calls shown in Figure 53, the reported values for other emergency calls were zero between 2015 and 2017 and increased to 84 and 19 for 2018 and 2019, respectively. The reported number of nonemergency calls also rose, with zero in 2015, one in 2016, 10 in 2017, 21 in 2018 and 124 in 2019. These trends need to be further evaluated to determine if reporting changes contributed to the increase in 2018.

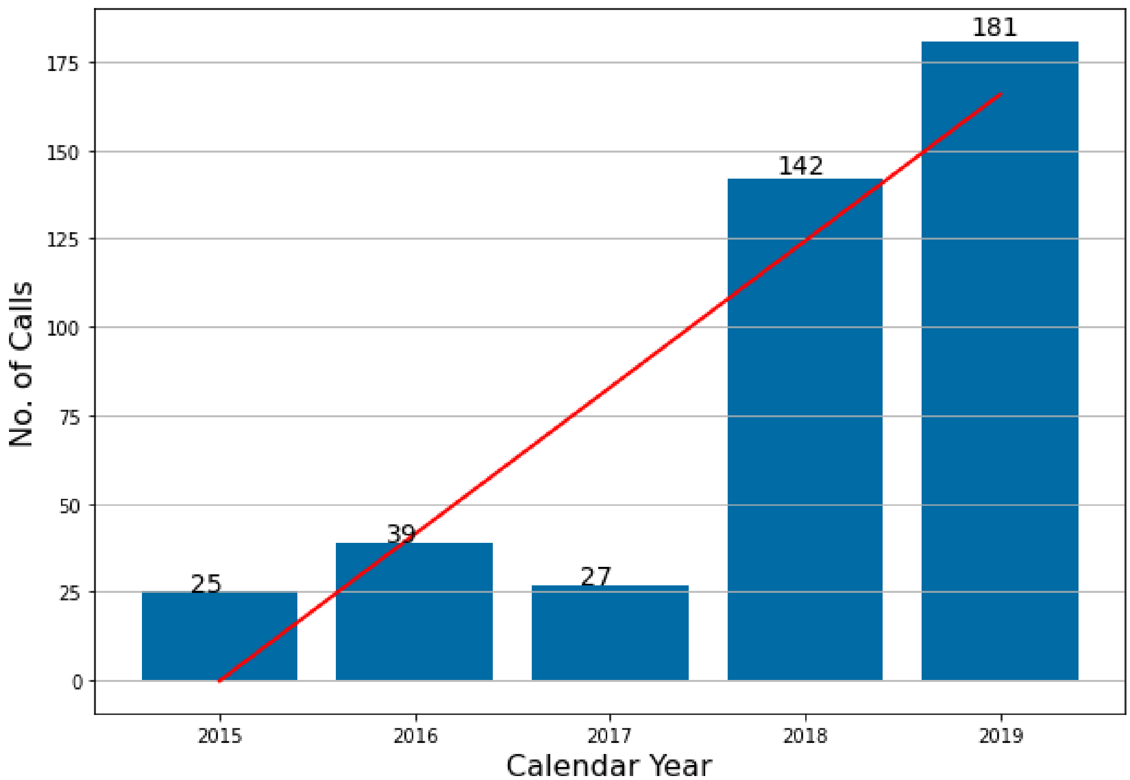

Figure 51: NREL Number of Fire Department Response Calls

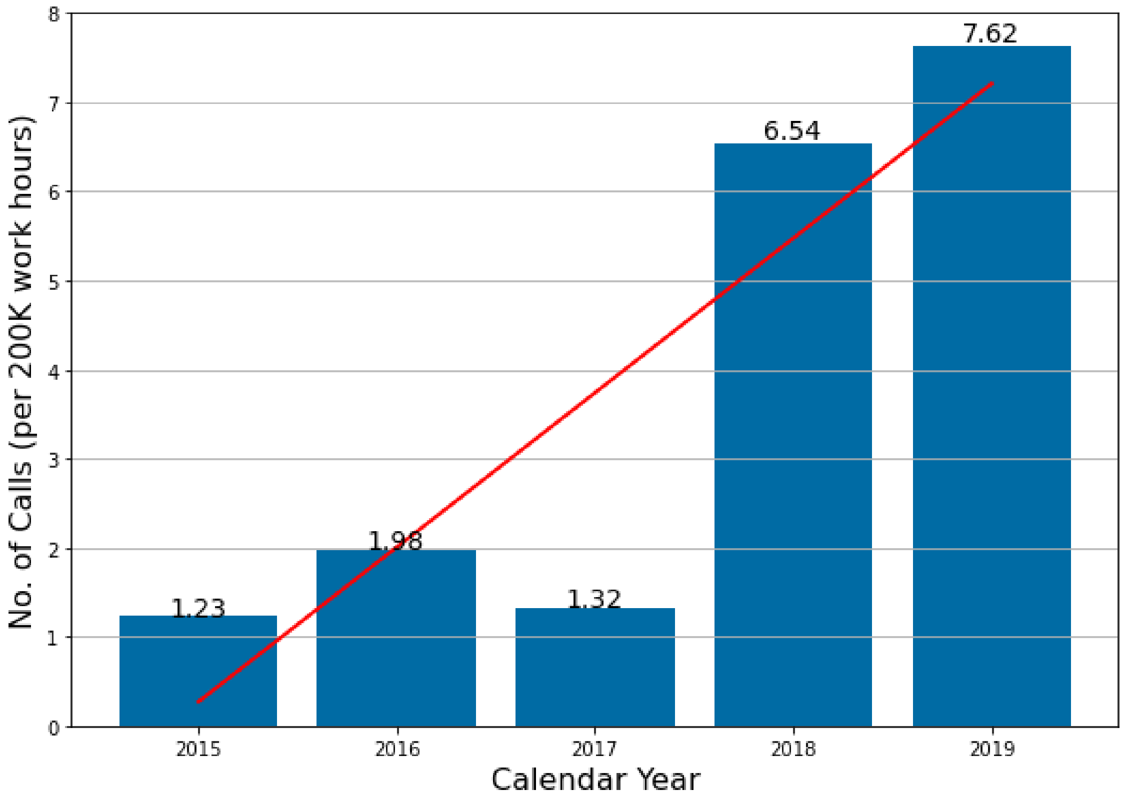

Figure 52: NREL Number of Fire Department Response Calls (per 200K site work hours)

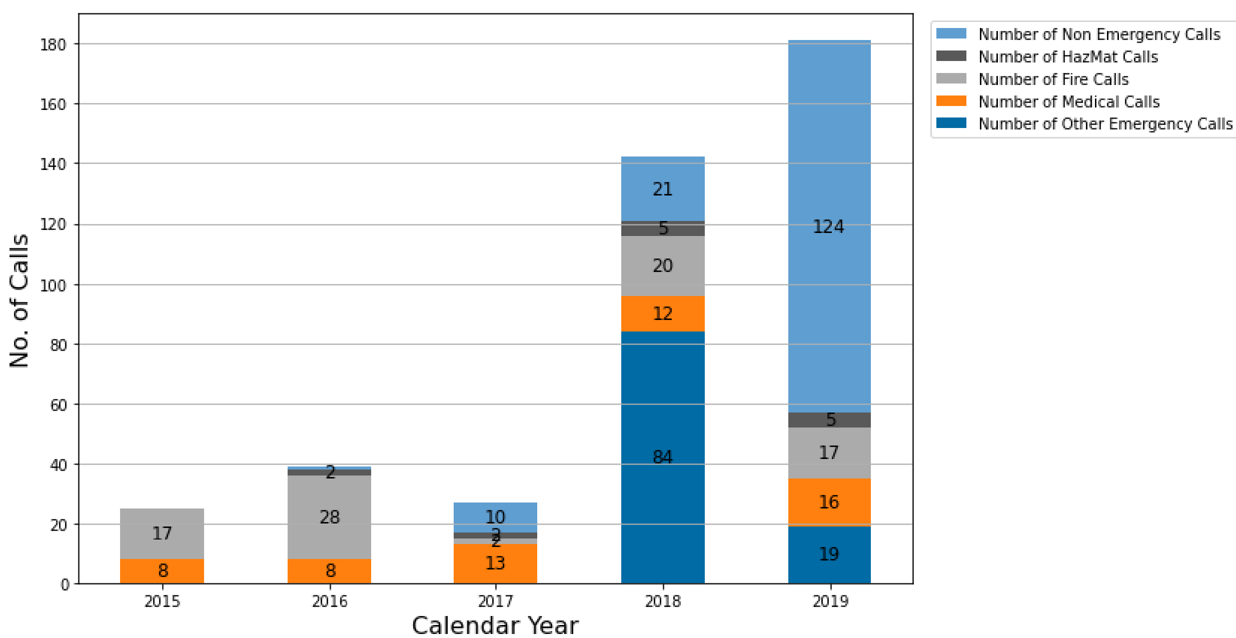

Figure 53: NREL Types of Fire Department Response Calls (Stacked Bar Plot)

4.2.4 Argonne National Laboratory (ANL)

Table 12 showed that the ANL site had one of the largest decreasing trend lines for the normalized total number of calls. Table 11 showed that the site had the third largest 2019 normalized total number of calls meaning that the site had a large number of calls per 200K work hours. Figure 54 show the trends for the total number of calls per CY and Figure 55 the normalized values. Both of these figures show that the trends are decreasing. The distribution of the types of calls, shown in Figure 56, shows consistent values across all the types of calls except for the number of medical calls. In 2019 zero medical calls were reported, most probably a missing value. This large decrease in the 2019 causes a decreasing linear trend.

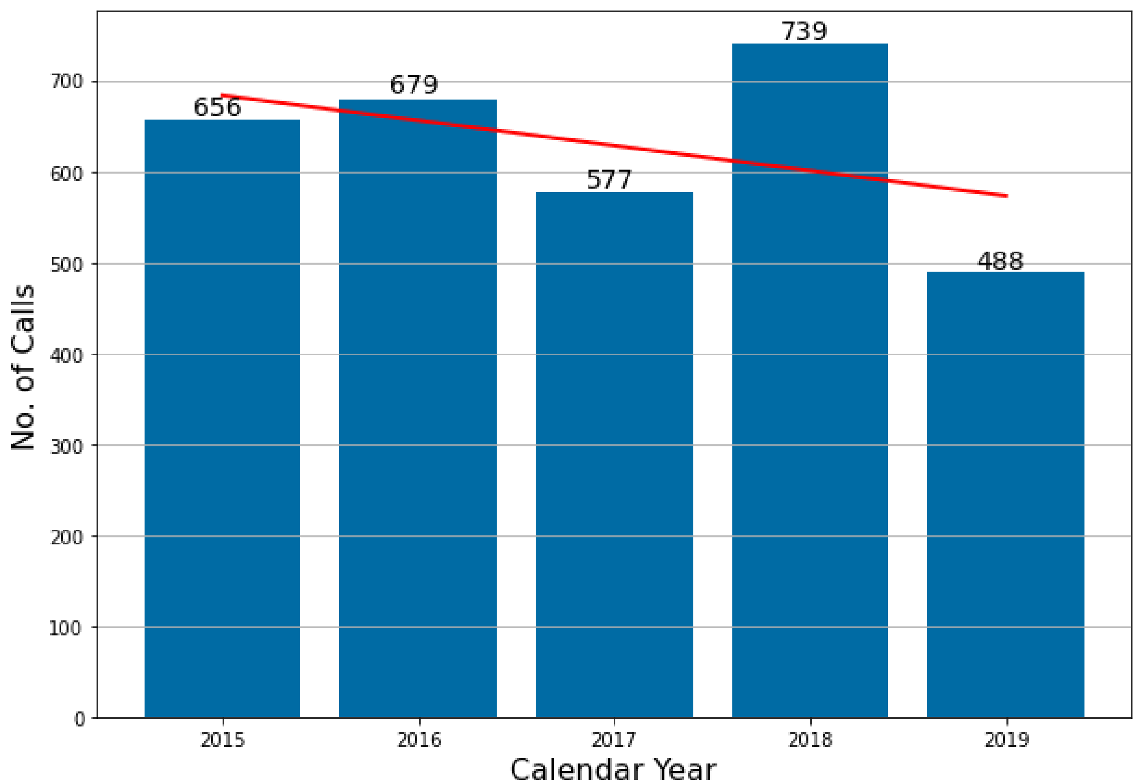

Figure 54: ANL Number of Fire Department Response Calls

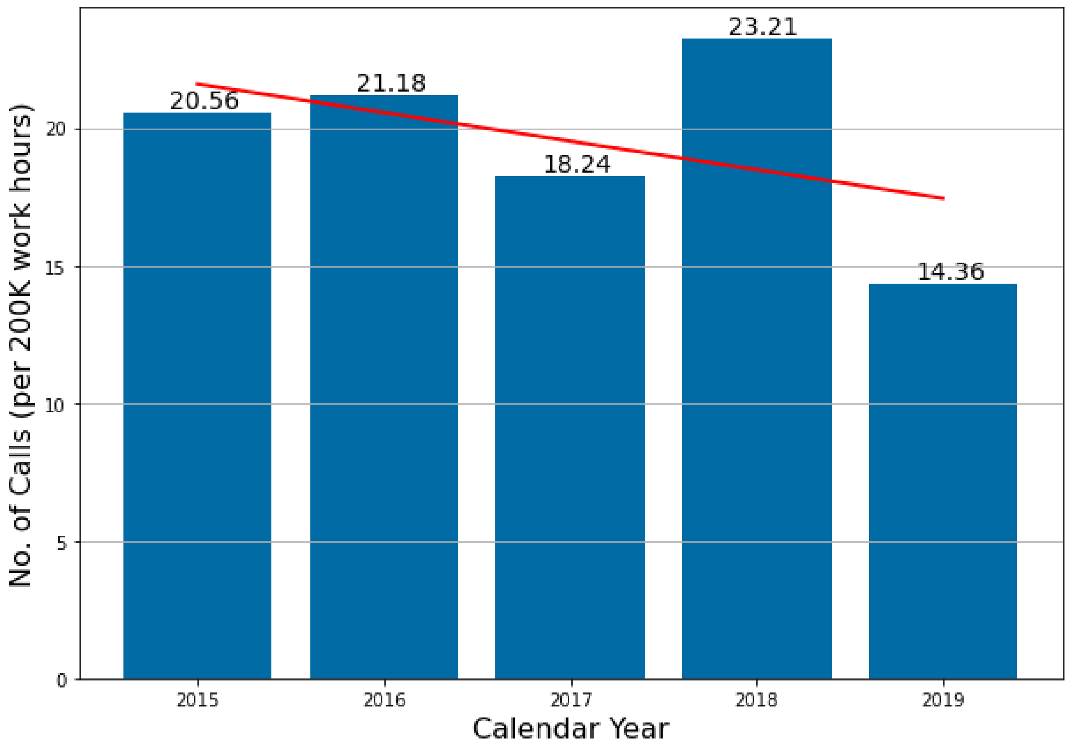

Figure 55: ANL Number of Fire Department Response Calls (per 200K site work hours per CY)

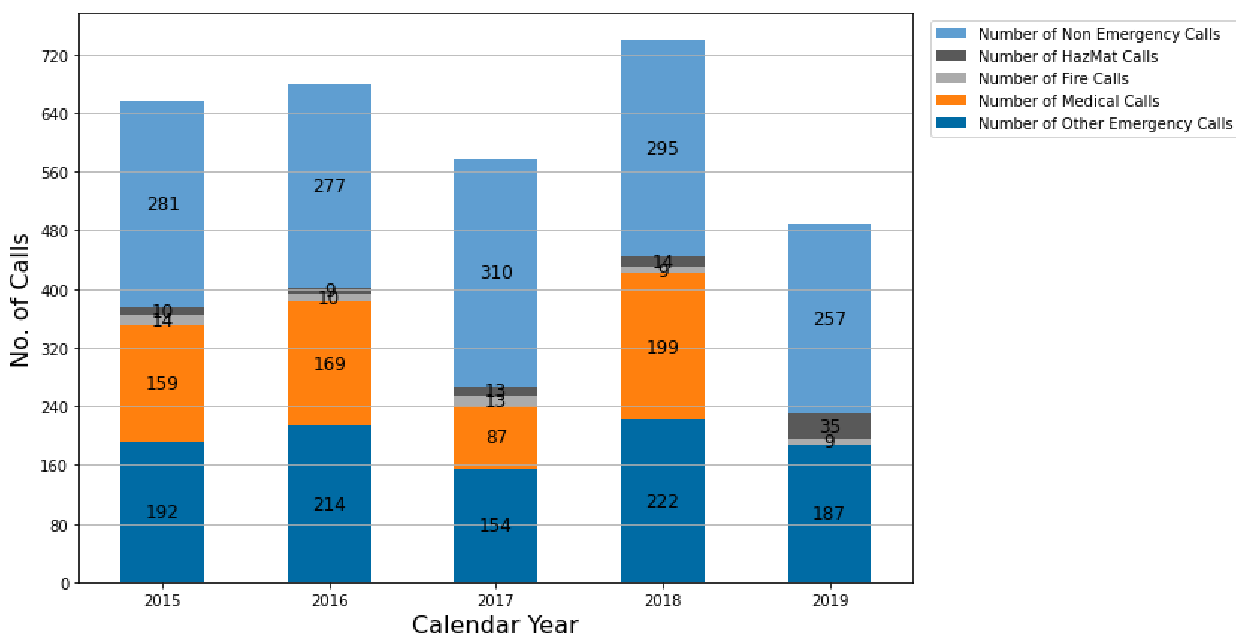

Figure 56: ANL Fire Department Response Calls by Type

5 Concluding Remarks

This report analyzed available data related to fire protection from multiple DOE databases. The data sources included the FPDB (monetary losses and fire department response calls), ORPS, and CAIRS for normalizing the results (i.e., per 200K work hours). Highlights of DOE wide trends:

- Trend for the number of reports in the DOE fire protection loss data is flat (see Figure 1), and the normalized trends are decreasing (see Figure 2).
- Trend for the total reported fire protection losses (non-normalized and normalized) are increasing (see Section 2.1).
- "Fire/Smoke (Brush)" type had the highest increasing trend by any measure (i.e., number of reports or monetary losses) (see Section 2.2).
- "Design/Material Related" cause had the highest increasing number of reports trend Section 2.3).
- "Other" cause had the highest increasing monetary losses trend (see Section 2.3).
- The trend for the number of fire protection reports greater than $50K is decreasing, however the trend for the monetary losses is increasing (see Section 2.4).
- Trend for ORPS fire protection reports (i.e., Keyword category 03-Fire Protection) is decreasing (see Section 3).
- The trends for fire department responses (non-normalized and normalized) are increasing (see Section 4), however there were observed limitations with the data, and the response calls trends are considered highly uncertain.

The analysis used machine learning to analyze the text of the descriptions in the fire protection loss reports (see Sections 2.5, 2.6, 2.7, and 2.8 for description and results of the analysis). The text analysis showed that the majority of the vegetation related fire reports (which is related to the Fire/Smoke (Brush) type fires) was caused by cigarettes, equipment failure (e.g., electrical arcing), downed power lines, and lightning strikes.

Other topics that resulted in identification of frequent topics within the reports description and resulted in high fire protection losses included cold weather-related events (i.e., which can result in frozen pipe bursts and building flooding), fire protection losses related to failure of equipment (e.g., transformers, fans, capacitors, and modulators), fire protection loss reports related to chemical reactions within "fume hoods", vehicle accidents and different types of microwaves (e.g., kitchen, analytical, etc.).

The analysis also evaluated fire department response data at four sites. This analysis identified potential limitations and challenges in the fire department response data potentially caused reporting differences from site to site. Improving reporting of the FPDB will ensure consistency and reliability in measuring trends to identify areas of improvements, and leading indicators that could enhance learning and safety across DOE sites. AU is working to improve data collection, data quality, ability to measure trends and identify areas of improvements and potential leading indicators that could enhance learning and safety across DOE. Improvements will be coordinated with the users of the DOE's FPDB as part of a separate initiative.