The National Fire Incident Reporting System and the Public Data Release File

Before extracting the U.S. Fire Administration's (USFA) National Fire Incident Reporting System (NFIRS) data from the annual Public Data Release (PDR) files, it is important to understand what NFIRS data is, what it is not, and how to use the data. This guide provides a brief overview of the NFIRS, its data, and its limitations in data analyses at the state and national levels, as well as analytic resources. A brief overview of the PDR files is also discussed.

What is NFIRS?

- NFIRS is a voluntary, all-incident based reporting system used by fire departments across the country to record detailed information about the incidents attended. NFIRS is based on a reporting standard that fire departments use to uniformly report on a wide range of their activities, from fire to Emergency Medical Services (EMS) to equipment involved in the response.
- NFIRS constitutes the world's largest, national, annual collection of incident information.
- All 50 states and the District of Columbia report incidents to NFIRS. State participation is voluntary, and each state specifies NFIRS reporting requirements for its fire departments. Regardless of state participation in the system, all fire departments are eligible to report to NFIRS. States that currently participate in NFIRS, may not have participated in previous years.
- The Department of Defense (DOD) and the Native American Tribal Authority also participate in NFIRS. Although DOD participates in NFIRS, their data is not released at the national level.
- Nationally, about 24,000 fire departments report in NFIRS each year. This includes over 30 fire departments with a population protected of more than 500,000. Participating fire departments report about 22,000,000 incidents and 1,000,000 fires each year.[1]

What NFIRS is not.

- NFIRS is <u>not</u> a survey based on a statistically selected sample.[2]
- NFIRS is <u>not</u> a complete census of reported incidents in the U.S.[3] Consequently, NFIRS does not capture all incidents and associated losses that occur across the U.S. each year. However, as previously noted, it collects about 1 million fire incidents and about 22 million

[1] USFA, NFIRS, https://www.usfa.fema.gov/data/nfirs/about/index.html.

[2] A *sample* is a subset of measurements selected from the population; a finite part of a statistical population whose properties are studied to gain information about the whole.

[3] A *census* is an official count or a complete enumeration of a population.

non-fire incidents each year from about 24,000 fire departments or 80 percent of the estimated 30,000 fire departments in the United States.[4]

What does all of this mean?

The raw data as reported to NFIRS can be used to:

- Identify issues that occur only rarely at the local and state levels, but may be detectable at the national level.
- Determine percentages of reported incidents by specific data elements.
- Derive reported loss per incident statistics.
- Determine leading causes and factors of specific types of fires.

Raw NFIRS data should not be used to represent total counts of incidents and associated losses (i.e., casualties and dollar loss) at the national or state levels. Note: if raw NFIRS data are used for analyses, it should be noted as such (i.e., results are based on "reported NFIRS data" or "data as reported to NFIRS").

Using raw NFIRS data alone at the national and state levels to represent total incident counts is not a proper use of the data because

- NFIRS participation is voluntary at the national level, and in most cases, the state level as well.
- NFIRS is not a complete census of reported incidents in the U.S. nor is it based on a statistically selected sample.

For these reasons, USFA computes national estimates.

What are National Estimates?

National estimates are estimates of the numbers of incidents associated with a subset of the data.[5] For example, where fire incidents are of interest, national estimates are estimates of the numbers of fires and losses (i.e., fires, deaths, injuries and dollar loss) associated with a subset of the fire data. High-level summarized national estimates of the numbers for fires, deaths, injuries and dollar loss are based on NFPA's annual Survey of Fire Departments for U.S. Fire Experience. With the exception of the NFPA estimates for total fires, structure (i.e., residential and nonresidential) fires,

[4] For 2015, the National Fire Protection Association (NFPA) estimated that there were 29,727 fire departments in the U.S. Source: NFPA, U.S. Fire Department Profile 2015, http://www.nfpa.org/research/reports-and-statistics/the-fire-service/administration/us-fire-department-profile, April 2017.

[5] An estimate is an approximation of a count or total.

vehicle, outside and other fires, all other estimates are scaled-up national estimates or percentages, not just the raw totals from NFIRS.

Because the NFIRS 5.0 data is not based on a statistically selected sample and does not represent a "complete" census of fire incidents, the raw counts of NFIRS data must be scaled up to national estimates. These estimates are based on a method of apportioning the NFPA estimates for total fires, structure fires, vehicle, outside and other fires. Generally speaking, the national estimates are derived by computing a percentage of fires, deaths, injuries or dollar loss in a particular NFIRS category and multiplying it by the corresponding total estimate from the NFPA annual survey.

The NFPA's Survey of Fire Departments for U.S. Fire Experience is based on a stratified random sample of U.S. fire departments. The sample of departments is stratified by size of community protected, and ratio estimation methodology is used to develop national-level summary estimates on fire loss statistics (the total numbers of reported fires, fire deaths, fire injuries and direct dollar loss) as well as summary estimates of fires and losses by major incident types (i.e., structure, vehicle, outside and other). For more detailed information on NFPA's survey methodology, refer to NFPA's report on fire loss in the U.S.: http://www.nfpa.org/news-and-research/fire-statistics-and-reports/fire-statistics/fires-in-the-us/overall-fire-problem/fire-loss-in-the-united-states.
National estimates are based on "The National Estimates Approach to U.S. Fire Statistics" by Hall and Harwood: http://www.nfpa.org/news-and-research/fire-statistics-and-reports/fire-statistics/how-nfpa-estimates-fires-and-fire-losses.

Is the data representative?

The percentage of fire departments participating in NFIRS varies from state to state, with some states not participating at all in some years. To the best that USFA can determine, the distribution of participants is reasonably representative of the entire nation, even though the collection of data is not based on a random sample. The dataset is so large and reasonably distributed geographically and by size of community that it is used as input to developing national estimates.

Moreover, most of the NFIRS data exhibit stability from one year to the next, without radical changes. Results based on the full dataset are generally similar to those based on part of the data, another indication of data reliability.

It is important to note that the USFA, along with other federal agencies, does not use NFIRS data to derive state-level estimates. NFIRS data is used to produce estimates at the national level.

What is the quality of the data in NFIRS?

Data quality is an area of great importance. Three criteria are used in the monitoring of the data in NFIRS during the year: (1) the data are complete, (2) the data are accurate, and (3) the data are current. These criteria are monitored by creating reports from the database that show the number of reporting fire departments, the number of incidents by state, the number of invalid incidents, and the number of unreleased incidents. The USFA provides the reports to the state NFIRS program managers and works with them to resolve any data issues. Technical assistance (e.g.,

telephone support or site visits) is provided to states to help address any data quality and data reporting needs.

Audits of the data are performed during the year to identify any inconsistencies. The audits focus on three criteria: gaps in reporting, critical errors in the data, and outliers in the data. In particular, USFA works closely with states to monitor the quality of data coming from third party vendor software. Each state is responsible for enforcing that the NFIRS third party software sold by vendors in their state is compliant with NFIRS standards. USFA assists states in monitoring vendor data quality issues or contacts vendors directly to discuss an issue at a state's request. Other data quality issues are questionable high dollar-loss incidents and questionable high numbers of fire deaths. Annually, USFA staff queries the database for questionable values (i.e., outliers) and verifies the values with state-level NFIRS program managers and local-level NFIRS program managers. The data quality steps are important to ensure that the data meet USFA's three criteria.

USFA published the report "Review and Assessment of Data Quality in the National Fire Incident Reporting System," (May 2017). This document covered a review of the system, the many robust data quality checks and mechanisms which are an integral part of NFIRS, and an assessment of the data quality both at the state level and at the data element level. The data element assessment focused on the most common data elements used in NFIRS data analyses. NFIRS data from the three most recent years available at the time of the report's development (2009 to 2011) were reviewed. Additionally, a section drawn from published NFPA documents covering the NFPA survey methodology was also included. The "Review and Assessment of Data Quality in the National Fire Incident Reporting System" document is available at: https://www.usfa.fema.gov/downloads/pdf/publications/nfirs_data_quality_report.pdf.

Unknown entries, incomplete loss reporting and unreported data are also important considerations when assessing NFIRS data quality. These topics are discussed in more detail in USFA's *Fire in the United States* 2006-2015 (19th Edition) report available at: https://www.usfa.fema.gov/downloads/pdf/publications/fius19th.pdf.

NFIRS Public Data Release Files

NFIRS 5.0 provides 11 modules that recognize the increasingly diverse activities of fire departments today. These modules, together, contain 567 data elements or fields.

The Basic Module is the main module, which is completed for every incident. The other modules are filled out, when appropriate, to provide additional information on an incident. All 11 modules are listed in the following table:

Module	Description
Basic Module	General information for each incident
Fire Module	Fire incident information
Structure Fire Module	Information on structure fires
Civilian Fire Casualty Module	Fire-related injuries or deaths to civilians
Fire Service Casualty Module	Injuries or deaths to firefighters
EMS Module	Medical incidents
Hazardous Materials Module	Hazardous materials incidents
Wildland Fire Module	Wildland or vegetation fires
Apparatus/Resources Module	Apparatus-specific information
Personnel Module	Personnel associated with apparatus
Arson Module	Intentionally-set fire information

Data from the modules are grouped together each calendar year to create the PDR files in delimited text (.txt) format, which are then released annually into the public domain. For NFIRS data submitted prior to 2012, the PDR files were released in dBASE (.dbf) format. The Apparatus/ Resources and Personnel Modules are excluded from the PDR because they are intended for local fire department use, and the PDR dataset's main utility is intended for national analyses. The PDR files consist of a subset of the data fields contained within the NFIRS national production database. For example, data elements with sensitive or identifying information are removed, as are data elements that are wholly used for maintenance or production purposes. The data structure of the PDR files has been considerably simplified from the production database's schema for ease of use.

The PDR files from 2004 to 2013 only include fire and hazmat incidents and their related data tables (available on CD). Prior to 2004, all incidents were included in the PDR files. Beginning with the 2014 NFIRS data, both the fire and hazmat incident PDR file (CD) and the full, all-incident PDR file (DVD) are available upon request from USFA's National Fire Data Center. The 2016 NFIRS PDR files are the most current data available.

In its basic form, the NFIRS PDR files have a relational data structure where data from each incident module is represented by a row in a data table. The primary tables (basic incident and incident address) contain most of the Basic Module data. There is exactly one record in the basic incident table for every incident reported to NFIRS. All other modules, represented by data tables with similar names (fire incident, civilian casualties, etc.), have records that are linked to the basic incident table through unique incident identification key fields (e.g., STATE, FDID,

INC_DATE, INC_NO and EXP_NO). Some module data are split across several tables (e.g., basic incident, incident address, and basic aid tables); one table (fire incident) combines data from two modules (i.e., Fire Module and Structure Fire Module). Some tables, such as fire incident, will only have one record for each relevant incident in the basic incident table, while tables such as civilian casualty may have several records linked to a single incident in the case where multiple injuries and/or deaths occur in the same incident.

As previously noted, state participation is voluntary, and each state specifies NFIRS reporting requirements for its fire departments. States have the flexibility to adapt their state reporting systems to their specific needs. As a result, the design of a state's data collection system varies from state to state. NFIRS 5.0 was designed so that data from state systems can be converted to a single format that is used at the national level to aggregate and store NFIRS data.

It is important to note that NFIRS underwent a substantial transition in 1999 to the Version 5.0 system. The system continued to accept legacy data (Version 4.1), however, since Jan. 1, 2009, NFIRS 4.1 data have no longer been accepted by the system. Starting with 1999, all Version 4.1 data in the system, regardless of the entry mechanism, are in NFIRS 5.0 format; non-NFIRS 5.0 data were converted to the 5.0 format.

The PDR files can be queried using SQL, ACCESS, or statistical software packages such as SAS, SPSS, or R. EXCEL is not recommended as it does not have the capability to support very large datasets and will result in the truncation of data rows that exceed 1 million rows.

The NFIRS PDR files include detailed documentation regarding the NFIRS data. Before analyzing the NFIRS data, USFA staff highly recommends that the documentation is reviewed to acquire a better understanding of the data and how to perform data analyses.

NFIRS Analytic Resources

Several resources are available that provide more detailed documentation on using the NFIRS and the NFIRS data. The "National Fire Incident Reporting System Complete Reference Guide" (https://www.usfa.fema.gov/data/nfirs/support/documentation.html) provides both instructions for reporting data to the NFIRS and an understanding of the data elements collected by the system. It also serves as a reference for coding the data. This document is also available on the PDR file.

The document "National Fire Incident Reporting System Version 5.0 Fire Data Analysis Guidelines and Issues" discusses analytic considerations and methods of analyzing fire-incident data using the NFIRS data. Topics include the NFIRS 5.0 data structure, general quality assurance issues, and definitions and parameters of common fire analyses (e.g., residential building fires or casualties), including the methodology for determining structure fire causes. The methods, techniques and considerations discussed are those used by the USFA analysts, and they do not necessarily reflect methods, techniques and considerations used by fire data analysts from other agencies and organizations. NFIRS data partners may (and do) employ their own methods for analyzing the data and may make differing assumptions when encountering data issues (http://www.usfa.fema.gov/downloads/pdf/nfirs/nfirs_data_analysis_guidelines_issues.pdf). This document is also available on the PDR.

As previously noted, "The National Estimates Approach to U.S. Fire Statistics" is the original methodology for creating estimates of the U.S. fire problem using the NFPA's annual Survey of Fire Departments for U.S. Fire Experience and the NFIRS data. The authors present a detailed consensus procedure for such calculations and the supporting rationale. This document is available at: http://www.nfpa.org/news-and-research/fire-statistics-and-reports/fire-statistics/how-nfpa-estimates-fires-and-fire-losses. "National Estimates Methodology for Building Fires and Losses" is the USFA's application of the national estimates approach to building fires and fire losses. It details USFA's current fire data estimation methodology for all building fires (i.e., residential and nonresidential) and associated losses (http://www.usfa.fema.gov/downloads/pdf/statistics/national_estimate_methodology.pdf).

The USFA's "Fire in the United States 2006-2015" is a statistical portrait of the national fire problem and provides an in-depth discussion of the data sources and the methodologies used to incorporate these data into fire analyses (https://www.usfa.fema.gov/downloads/pdf/publications/fius19th.pdf). Lastly, the "Fire Data Analysis Handbook" is a resource for those unfamiliar with basic data analysis techniques and their applicability to fire-data based analyses (http://www.usfa.fema.gov/downloads/pdf/publications/fa-266.pdf).