

Summarizing contemporary large-fire occurrence for land and resource management planning

Prepared by:

Joe H. Scott
Pyrologix LLC
Missoula, MT
joe.scott@pyrologix.com

Prepared for:

Elizabeth Reinhardt
Assistant Director, Fuel and Fire Ecology
USDA Forest Service, Washington Office
ereinhardt@fs.fed.us

September 9, 2014

version 0.1

DRAFT

Executive summary

When it comes to wildfire management planning, the recent past may be a guide to the near future. As such, summarizing contemporary historical wildfire occurrence for a designated portion of the landscape (a fire occurrence area, or FOA) is a valuable analysis that contributes to fire management planning. This whitepaper describes one such wildfire occurrence analysis process.

The preparatory steps in the process are to delineate the fire occurrence area (FOA) for the project and then to obtain or create a fire occurrence database (FOD) for the FOA. The analysis process produces nine results:

1. Lorenz curve
2. Empirical cumulative distribution function
3. Large-fire size threshold
4. Large-fire season chart
5. Seasonal chart of fire occurrence
6. Contemporary large-fire occurrence rates
7. Large-fire ignition density grid
8. Large-fire day distribution
9. Logistic regression coefficients

Eight of the nine can be generated with only the FOD for the FOA. The last result—logistic regression coefficients—also requires a database of daily values of the Energy Release Component (ERC) of the National Fire Danger Rating System (NFDRS) for fuel model G.

These nine analysis results are useful for planning for future large fires in the FOA, and for parameterizing and calibrating a stochastic large-fire simulator like FSim.

1. Introduction

Analysis of contemporary wildfire occurrence has been a key component of wildfire management for almost a century (Short 2014). Today, contemporary large-fire occurrence data are used to parameterize and calibrate wildfire simulations systems, like FSim (Finney and others 2011), that are used to assess wildfire risk and plan fuel management activities. In this document, fire occurrence refers to an instance (occurrence) of a wildfire. A wildfire occurrence database compiles important characteristics of each wildfire, including its date of discovery, start location, and final size.

This document addresses *contemporary* wildfire occurrence, meaning that it refers to wildfire occurrence of the recent past, but still closely connected to today. Contemporary wildfire occurrence can be generally taken to mean the wildfire occurrence of the last 20-30 years. The fire occurrence of 30-50 years ago does not necessarily apply to today because so many factors—fire climate, fuel continuity and volume, fire suppression strategies, fire management practices, and so on. Contemporary wildfire occurrence is also much different than presettlement fire history. Presettlement fire history refers to the frequency of fire and mean area burned by wildfire during the time before widespread US settlement by Europeans, usually taken to be the beginning of the 20th century. During the presettlement era, factors affecting wildfire occurrence were drastically different from today, especially the lack of organized wildland fire protection during that time.

The analyses described in this document are designed to assess the occurrence of the relatively few large wildfires that account for the overwhelming majority of area burned (Strauss 1989). These analyses are not necessarily sufficient for planning or implementing preparedness and pre-suppression activities, which are sensitive to *all* fire occurrences, not just those that become large.

1.1 Datasets needed

Two datasets are needed for the large-fire occurrence analysis described in this report. The majority of the analysis (seven out of eight steps) can be completed with just one of these datasets, a spatial fire-occurrence database (FOD) such as that produced by Short (2014) for the continental US. Only three attributes are required for each wildfire in the FOD:

- the start location
- start (or discovery) date
- final size

Other attributes, such as cause or jurisdiction, may be helpful for a more detailed analysis for a specific purpose. Although the analysis focuses on large wildfires, a complete "all fires" dataset is useful for identifying the threshold "large fire" size.

The eighth and final analysis step, which relates large-fire occurrence to ERC-G, can only be completed with the addition of a tabular database of daily values of the Energy Release Component (ERC) of the National Fire Danger Rating System (NFDRS). The ERC must be calculated specifically for NFDRS fuel model G (ERC-G). The ERC-G values must be available for the same time period as the FOD, for every calendar day of the year. ERC-G values are typically calculated in FireFamilyPlus from data measured at a RAWS station. Spatial datasets of daily values of ERC-G for the CONUS have recently become

available (Jolly, personal communication). Such a dataset can be used to produce an ERC-G dataset for any location.

1.2 Results produced

The large-fire occurrence analysis produces nine individual results. Each result is a graph, table, value, or set of coefficients for a defined fire occurrence area (FOA). A FOA is a geographic area within which contemporary large-fire occurrence characteristics can be assumed to be uniform.

1. **Lorenz curve**—The Lorenz curve for wildfire occurrence plots the cumulative fraction of wildfires, from smallest to largest, on the X-axis against the cumulative fraction of area burned on the Y-axis.
2. **Empirical cumulative distribution function**—The empirical cumulative distribution function (ECDF) of wildfire occurrence is similar to the Lorenz curve, but the X-axis is final fire size (again sorted from smallest to largest) and the Y-axis is cumulative number per year or fraction of fires less than or equal to that size.
3. **Large-fire size threshold**—A summary of large-fire occurrence needs a definition of "large fire". Several approaches are available to identify the size threshold for defining a large fire.
4. **Large-fire season chart**—This chart plots the cumulative area burned (by large fires) against Julian day of discovery. It is useful for identifying the large-fire season, defined as the period of the calendar year during which large-fires ignite and eventually burn the majority of large-fire area.
5. **Seasonal chart of fire occurrence**—This chart plots the final size of each wildfire in the historical record against its Julian day of discovery.
6. **Contemporary large-fire occurrence rates**—This is a simple tabulation of the mean annual number, size, and area burned in large fires, normalized to a per-million-acres basis to facilitate comparison among landscapes of differing size. The values in the table are useful for calibrating large-fire occurrence and growth simulations made with FSim.
7. **Large-fire ignition density grid**—Large-fires ignitions are assumed to be a stochastic process, but it is not necessary to assume the likelihood of a large-fire ignition is spatially uniform. The large-fire ignition density grid (IDG) indicates the spatial likelihood of large-fire ignitions across the landscape.
8. **Large-fire day distribution**—A large-fire day is a day in which at least one large fire was discovered. A single large-fire on a large-fire day is the most common outcome. This frequency distribution of the number of large fires per large-fire day is tabulated to indicate how often multiple fires start on a single day.
9. **Logistic regression coefficients**—After coupling the tabular large-fire occurrence dataset with the ERC-G dataset, logistic regression coefficients can be generated for the probability of a large-fire day in relation to ERC-G.

1.3 Interpreted example

The contemporary large-fire occurrence analysis process will be illustrated using the Island Park Sustainable Forest Community (IPSFC) project on the Caribou-Targhee National Forest, encompassing the community of Island Park, Idaho (**Figure 1**). The IPSFC project required the use of the FSim fire occurrence and growth simulator. FSim requires several large-fire occurrence characteristics as input and for calibration.

2. Analysis process

2.1 Delineate the fire occurrence area

A fire occurrence area (FOA) is a geographic area for which contemporary fire occurrence characteristics can be assumed to be relatively uniform. More simply, it is the portion of the landscape for which contemporary fire occurrence will be summarized. Most FOAs are contiguous polygons, but that is not a requirement. For example, a particular vegetation type within a region could be used as a FOA, even if the vegetation type is not in one contiguous block, and even if the vegetation type is mapped as raster data.

The FOA size and location is determined, in part, by the land unit being assessed, and may include a significant buffer around the assessment unit, especially if stochastic fire growth modeling will be used. Examples of assessment units include a National Forest or National Park, or a Fire Planning Unit (FPU). The FOA

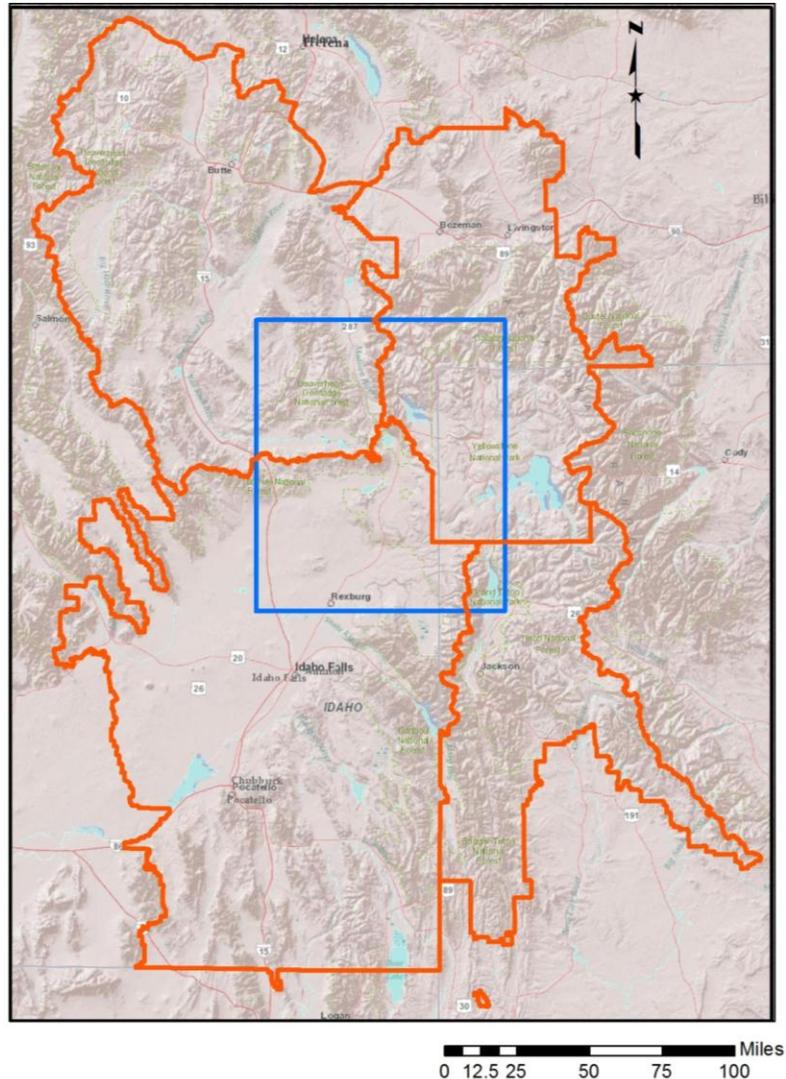


Figure 1—The 5.4-million acre fire modeling landscape (LCP) extent (blue rectangle) for the Island Park Sustainable Forest Community (IPSFC) project, located in the community of Island Park, Idaho, USA, includes the IPSFC project area plus a large surrounding buffer. The four Fire Planning Units (FPUs) that occur on the IPSFC LCP, totaling 30.8 million acres, are outlined in red. Clockwise from the northwest, the FPUs are: Headwaters, Greater Yellowstone Area North, Western Wyoming, and Eastern Idaho.

for a Forest or Park, including a buffer around them, can range from a few million up to several tens of millions of acres. Fire Planning Units range in size, roughly, from 1 to 45 million ac. If necessary, several adjacent FPU's can be combined into a single large FOA, as we do in our interpreted example.

For the IPSFC project, we will initially examine two different FOAs. The smaller FOA is the fire modeling landscape for the project, which encompasses 5.4 million acres including a buffer around the project area (**Figure 1**). The larger FOA is the four FPU's combined (**Figure 1**), encompassing 30.8 million acres.

2.2 Develop a fire occurrence database

After delineation of one or more FOAs for the project, the next step is to obtain or create a spatial FOD specifically for the delineated FOA(s). Recall that the only required attributes for this analysis are the start location, start (or discovery) date, and final fire size. It is important that this FOD cover all ownerships across the entire FOA. Some land management units may have such a FOD available locally. A spatial FOD covering the entire US (Short 2014), developed for the Fire Program Analysis (FPA) project, is an excellent source of contemporary fire occurrence data analyses like those presented here. The current edition of the FPA FOD spans 21 years, from 1991 to 2012, and is publicly available from the US Forest Service Research Data Archive (<http://www.fs.usda.gov/rds/archive/Product/RDS-2013-0009.2/>). Like many fire occurrence data sources, it is more reliable for large fires than for small fires (largely due to voluntary, and thus generally incomplete, reporting of fires on wildland under the jurisdiction of a local fire protection organization).

Because we wish to characterize *contemporary* fire occurrence, the time period for the FOD should ideally be the most recent 20-30 years. Wildfire occurrence data beyond about 30 years are not necessarily desirable, even if reliable, because fire occurrence during that time period may not represent the present-day fire occurrence rate.

The reliability of fire occurrence data for small fires is questionable in most datasets. However, the large-fire occurrence data are much more reliable. With the exception of the use of the all-fires Lorenz curve to identify a large-fire size threshold, small-fire data are not used in this analysis except to provide ancillary information, so any inaccuracy or incompleteness in those data is of little consequence.

For the IPSFC project, two FODs were created by selecting records from the FPA FOD that fell within the boundary of the IPSFC FOAs as shown in **Figure 2**. The IPSFC project-area FOD, spanning just 5.4 million acres, contains 1,166 fire records, of which 69 were greater than 100 ha. Because there were so few large-fire records, a second FOD was also generated for the combined four FPUs that intersect the IPSFC project area. Summary characteristics of these FODs are shown in Table 1. Because of the relative lack of records in the IPSFC fire modeling landscape, the combined four-FPU FOA will be used in the IPSFC interpreted example. Short (2014) provides information about the completeness of fire occurrence records by state and time period.

2.3 Summarize contemporary historical fire occurrence

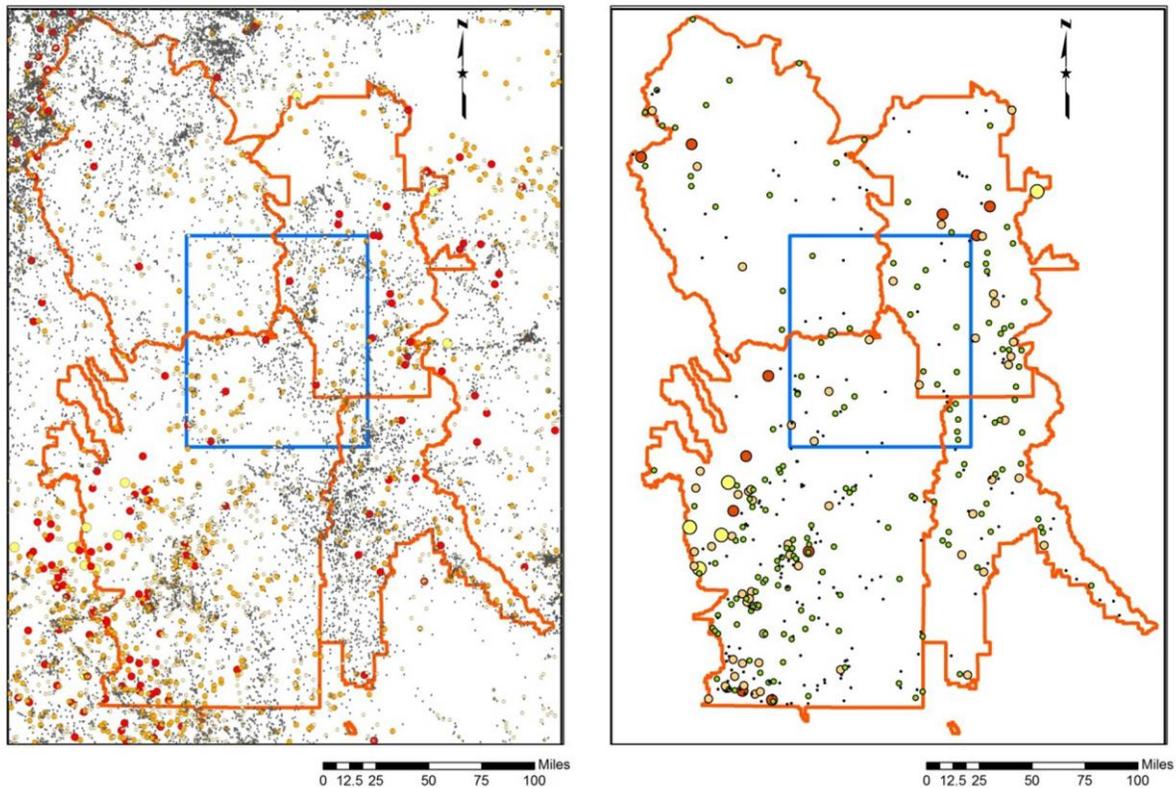


Figure 2—Graphical display of the IPSFC fire occurrence database (FOD) for all fires of any size (left) and for fires larger than 100 ha within the four Fire Planning Units that occur within the IPSFC landscape (right).

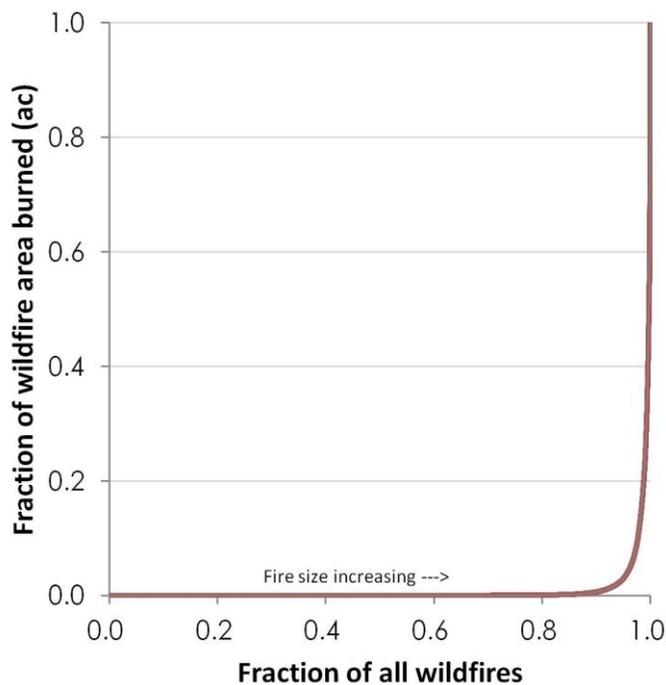


Figure 3. Lorenz curve for wildfire occurrence plots the cumulative fraction of number of wildfires on the X-axis (all fire sizes) against the cumulative acres burned (Y-axis). The dataset must be sorted by fire size before calculating the cumulative fraction of wildfire area burned.

distribution. **Figure 3** is a Lorenz curve for the four-FPU FOD surrounding the IPSFC project area. From **Figure 3** we can see that the smallest 90 percent of wildfires are responsible for only a tiny fraction—less than one percent—of the total area burned. Examining the FOD itself reveals that the 90th percentile fire size is just 50 acres, further illustrating the unequal distribution of the number of fires and their area burned—large fires do all the work. There are a number of ways to identify a large-fire threshold from this curve; they are discussed in section 2.3.3 below.

2.3.2 Empirical cumulative distribution function (ECDF)

A chart closely related to the Lorenz curve is the empirical cumulative distribution function (ECDF) of fire sizes in the FOD. The ECDF is a plot of final fire size (X-axis) against the cumulative probability of a fire of that size or smaller occurring in the dataset (**Figure 4**). Just like for the Lorenz curve, the dataset

This section outlines the process to summarize contemporary large-fire occurrence within the 30.8 million acre combined four-FPU area.

2.3.1 Lorenz curve

The first step in the analysis process is to build an annualized distribution of cumulative number of fires vs. cumulative acres burned—called a Lorenz curve—for the FOA. This curve should include all fire sizes occurring in the FOA, not just large fires. This curve will be used to identify a large-fire threshold (see section 2.3.2) and is useful for later comparison with FSIM results.

A Lorenz curve is created by sorting the FOD by fire size, smallest first, and then calculating, for each fire record, the cumulative fraction of number of fires and area burned accounted for at its point within the

Table 1—Summary characteristics of alternative fire occurrence datasets for potential use in the Island Park Sustainable Forest Community project.

FOD	Area (million ac)	Total number of fires	Large-fire size (ac)	Number of large fires	Number of years with no large fires
IPSFC	5.4	1,166	115	69	6
Four-FPUs	30.8	7,586	294	399	1

must be sorted by fire size before calculating the cumulative probability.

The shape of the ECDF is sensitive to whether all fires or just large fires are included in the plot.

Figure 4 plots curves for both all fires and for fires ≥ 100 ha (247 ac). The FSim large-fire simulation system produces valid fire-size results only for large fires, not small fires, so if the ECDF is to be used for comparison with FSim results, it should show the distribution for large fires in addition to the all-fires distribution.

From **Figure 4** we can see, again, that more than 97% of all fires are relatively small (less than 1000 ac) compared to the largest fires that occurred (up to 200,000 ac). Among the fires greater than 100 ha, 90 percent of them were smaller than 12,000 ac.

2.3.3 Large-fire size threshold

There is no single definition or size threshold to identify a "large" wildfire. Instead, there exist several alternative methods for identifying a large-fire size for the fire-size distribution for a FOA (determined with the FOD). Even though each method gives a different large-fire size threshold value, the differences among the methods in terms of large-fire area burned is relatively small; any of the methods can be used as long as the same method is used for all FOAs being compared. Likewise, the same threshold must also be used when calibrating a large-fire simulator like FSim.

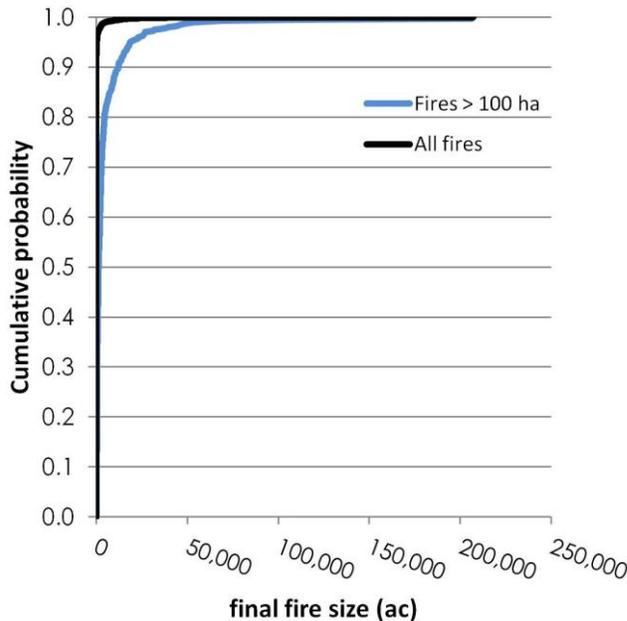


Figure 4. Empirical cumulative distribution functions for the four-FPU fire occurrence area. The ECDF plots the cumulative probability (percentile) or large-fire size against fire size.

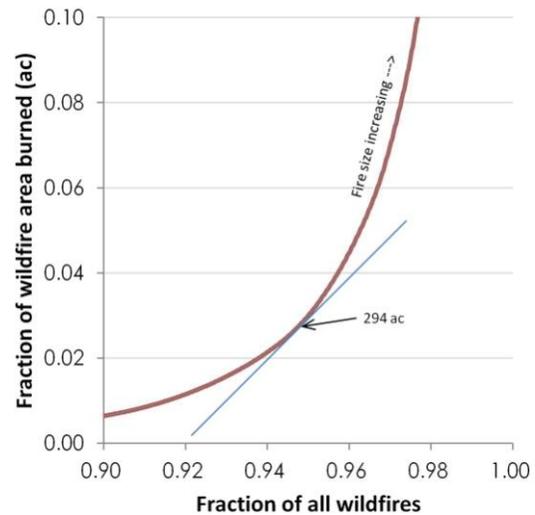


Figure 5. Lorenz curve (zoomed-in on the lower-right corner) illustrating the large-fire size threshold determined by identifying the fire size where the slope of the curve = 1.

Lorenz curve slope

One approach to identifying a large-fire size threshold is to find the fire size where the slope of the Lorenz curve equals 1 (**Figure 5**). That is, the size at which a one percent increase in cumulative fraction of wildfires equals a one percent increase in cumulative area burned.

For the fire-size distribution for the four-FPU FOD, encompassing 30.8 million acres, the fire size where the slope of the Lorenz curve = 1 is 294 ac. Fires smaller than this size account for less than 3 percent of the total area burned in this FOA.

Balanced fires-acres percentiles

Another approach to identifying a large-fire size threshold is to find the fire size where the cumulative percentage of wildfires is "balanced" by the cumulative percentage of area burned—that is, they sum to 100 percent. For example, the fire size (S) for which X percent of the fires smaller than S account for $100-X$ percent of the area burned. This balance point can be visualized graphically on the Lorenz chart by drawing a diagonal line from the upper-left to the lower-right corner of the chart (**Figure 6**). For this fire-size distribution for four Fire Planning Units, encompassing 30.8 million acres, the large-fire size threshold is 505 ac; 95.8 percent of the fires were smaller than 505 acres, but those fires accounted for only 4.2 percent of the total area burned.

Fixed percentage of total area burned

The third approach is to use a fixed percentage of the total area burned. For example, regardless of the shape of the fire-size distribution, the fire size that accounts for a given percentage of the total area burned (say 97 percent) could be identified from the Lorenz curve (**Figure 7**). For this fire-size distribution for the four-FPU FOD, the large-fire size that accounts for 97% of the total area burned is 335 ac.

Fixed size threshold

Finally, and most simply, a constant large-fire size (such as 300 ac or 100 ha) could be used on all FOAs, regardless of the shape of the distribution. This method cannot be visualized on the Lorenz curve, but its result (for 100 ha, or 247 ac) is shown in **Figure 8**.

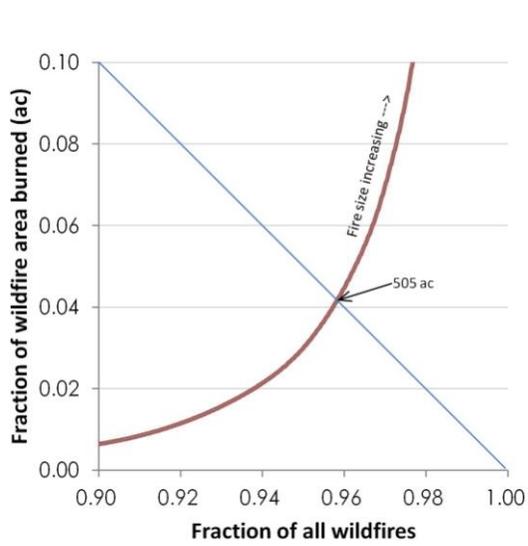


Figure 6. Lorenz curve (zoomed-in on the lower-right corner) illustrating the large-fire size threshold determined by identifying the fire size where fraction of all wildfires is balanced by the fraction of wildfire area burned.

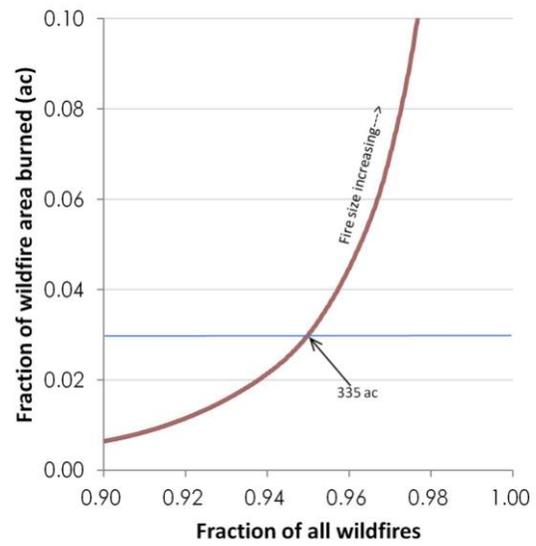


Figure 7. Lorenz curve (zoomed-in on the lower-right corner) illustrating the large-fire size threshold determined by identifying the fire size threshold that accounts for 97 percent of the area burned.

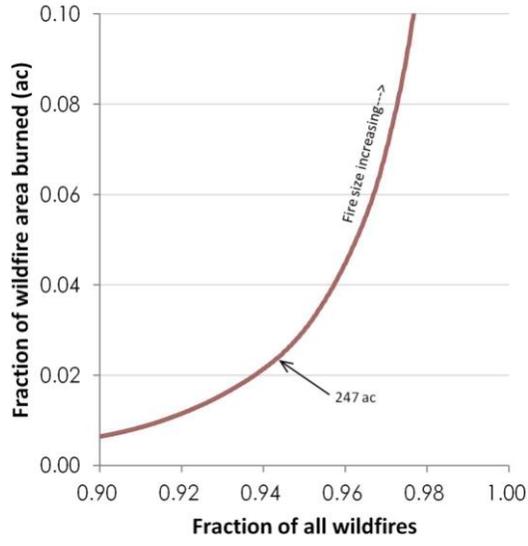


Figure 8. Lorenz curve (zoomed-in on the lower-right corner) illustrating where on the curve an arbitrary large-fire size threshold of 247 ac (100 ha) located.

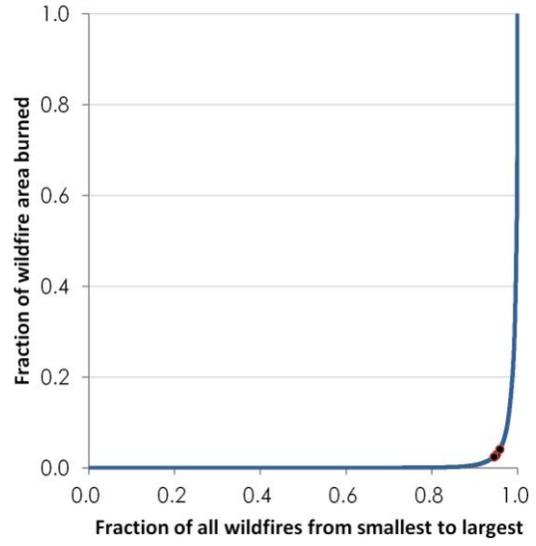


Figure 9. Lorenz curve showing all four alternatives for identifying a large-fire size.

Table 2. Alternative large-fire size thresholds and associated characteristics for the 30.8-million acre area covered by four FPUs. Even though the identified size thresholds vary by more than a factor of two, the associated total area burned varies by less than 2 percent.

	Large-fire size (ac)	percentage of fires larger than that size	percentage of area burned in fires larger than that size	Mean annual area burned in fires larger than that size
Lorenz slope = 1	294	5.3	97.3	109,166
Balanced	505	4.2	95.8	108,818
97th percentile	340	5.0	97.0	109,473
Constant size	247	5.5	97.5	107,543

Comparison of large-fire size methods

For the four-FPU FOD, these large-fire size thresholds determined by these four methods vary by more than a factor of two, from 247 acres to 505 acres. That range of threshold large-fire sizes makes little difference to the total large-fire area burned (Table 2), because most of the area burned is accounted for in fires much larger than this.

The four large-fire size thresholds are shown as points along the Lorenz curve in **Figure 9**. Again, all four methods adequately reflect the inflection point where fires larger than the threshold begin to have an outsized effect on area burned. In the remainder of this interpreted example we will use the large-fire threshold of 294 ac, which is size where the slope of the Lorenz curve = 1.

2.3.4 Large-fire season

It may be helpful to identify the large-fire season for a FOA. The large-fire season is defined generally as the time of year that accounts for the majority of large-fire activity. Large-fire activity can be defined by the numbers of large fires or by the area they burn. In this analysis, we identify the primary large-fire season by examining the cumulative distribution of large-fire area burned through the season based on discovery dates. The beginning of the large-fire season is identified as the date before which the large fires account for only 5 percent of the total. Likewise, then ending date of the fire season is identified as the date after which large fires account for another 5 percent of the total. That leaves 90 percent of the total within those dates. For the four-FPU FOD used for the IPSFC project, 5 percent of the cumulative area burned occurred in wildfires that started before Julian day 186 (July 6) and 5 percent occurred in

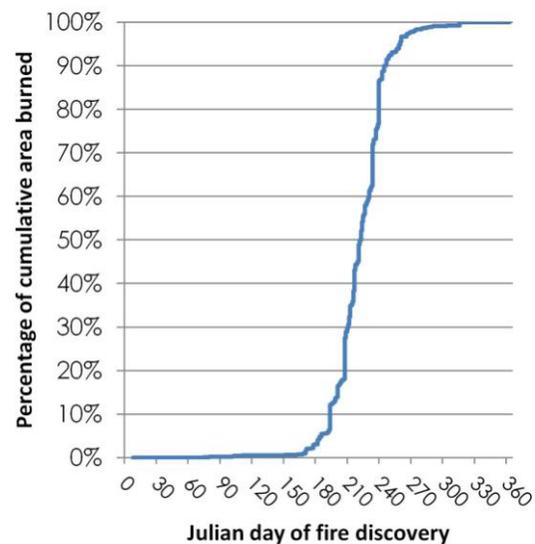


Figure 10. This large-fire season chart is a plot of cumulative percentage of large-fire area burned against the Julian Day of fire discovery.

wildfires that started after Julian day 260 (September 18), suggesting a 74-day large-fire season (**Figure 10**).

Alternatively, the large-fire season could be identified using the numbers of fires rather than their area burned. For the four-FPU FOD, 5 percent of the large fires were discovered before June 11 and another 5 percent were discovered after October 3. Both of these large-fire season dates are plotted on the seasonal fire occurrence chart in the next section (**Figure 11**).

2.3.5 Seasonal chart of fire occurrence

A seasonal chart of wildfire occurrence is a scatterplot with Julian day of wildfire discovery on the X-axis and final fire size (logarithmic scale) on the Y-axis (**Figure 11**). For reference, vertical lines representing the primary large-fire season as calculated above are shown as gray lines. The large-fire size threshold used is shown as a horizontal black line. The width of this black line indicates the secondary large-fire season, based on the number of large fires rather than their area burned.

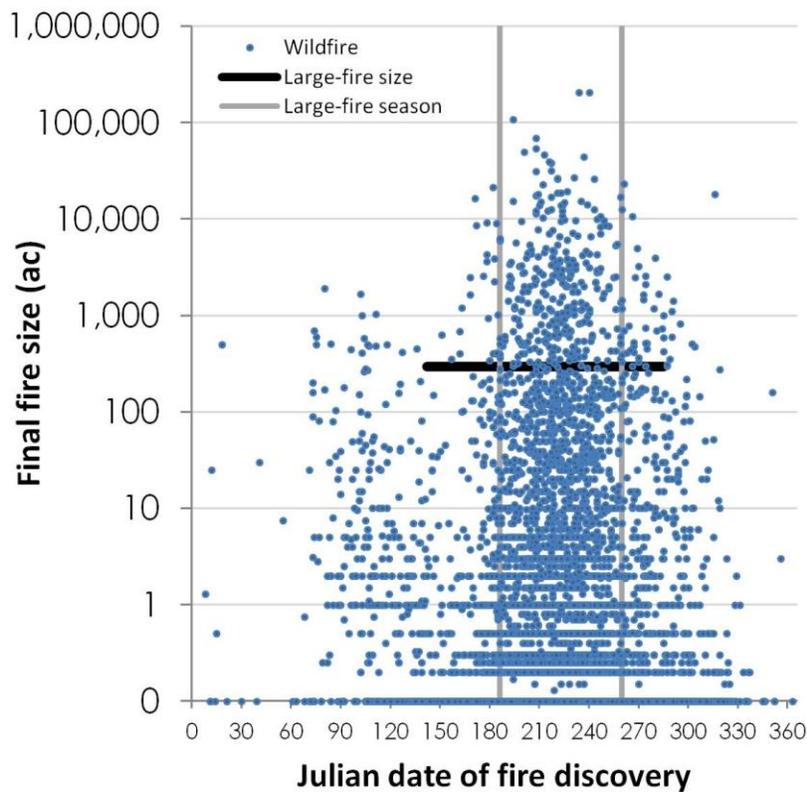


Figure 11. Seasonal chart of wildfire occurrence. Vertical gray reference lines indicate the primary large-fire season (capturing 90 percent of the large-fire area burned). The horizontal black reference line indicates the large-fire size threshold (294 acres); its width indicates the secondary large-fire season, capturing 90 percent of the large fires by number.

2.3.6 Contemporary large-fire occurrence summaries

In addition to the preceding charts, some simple summaries of contemporary large-fire occurrence rate and area burned are needed to parameterize and calibrate large-fire simulation models like FSim. The summaries begin with a tabulation of the number of large fires and total large-fire area burned for each year in the FOD (**Table 3**).

Several summary statistics can be calculated from this tabulation, including:

Table 3. Annual large-fire occurrence for the four-FPU FOD used for the IPSFC project.

Year (<i>y</i>)	<i>NumFires_y</i>	<i>AB_y</i>
1992	26	86,715
1993	0	0
1994	33	104,012
1995	13	19,942
1996	25	248,761
1997	1	2,218
1998	14	66,989
1999	20	90,426
2000	61	356,061
2001	27	64,366
2002	7	21,599
2003	32	133,564
2004	1	620
2005	12	46,460
2006	29	389,202
2007	45	229,061
2008	16	49,361
2009	7	26,551
2010	15	167,143
2011	15	80,573
SD	15.4	114,198
min	0	0
median	15.5	73,781
mean	20.0	109,181
max	61	389,202

- mean annual area burned
- mean annual burn probability
- mean annual number of large fires
- mean large-fire size

The mean annual area burned in the FOA is calculated as

$$\overline{AB} = \frac{\sum AB_y}{NumYears}$$

where AB_y is the large-fire area burned in year y , and $NumYears$ is the number of fire years in the database. For the four-FPU FOD, which spans a 20-year period from 1992 through 2011, a total of 2,183,623 acres burned in large fires. The calculation is therefore

$$\overline{AB} = \frac{2,183,623}{20} = 109,181 \frac{ac}{yr}$$

The mean annual BP is estimated by dividing the mean annual area burned by the size of the FOA, which in this case is 30.8 million acres.

$$\overline{BP} = \frac{109,181}{30,770,000} = 0.00355$$

Mean BP is useful for comparing FOAs of different size because it normalizes by the size of the FOA. The mean annual number of large fires ($NumFires$) is calculated as

$$\overline{NumFires} = \frac{\sum NumFires_y}{NumYears}$$

where $NumFires_y$ is the number of large fires in year y . For the four-FPU FOD, there were a total of 399 large fires over 20 years, so

$$\overline{NumFires} = \frac{399}{20} = 19.95 \frac{fires}{year}$$

Finally, the mean large-fire size can be calculated as

$$\frac{\sum AB_y}{\sum NumFires_y}$$

For the four-FPU FOD, the mean large-fire size is

$$\frac{109,181}{399} = 5,473 \text{ ac}$$

The mean large-fire size is a useful measure when calibrating FSim because it gives an indication of whether the simulated fires need to be, on average, larger or smaller in order to compare favorably with the historical large-fire sizes.

These summary measures can be plotted on a large-fire occurrence characteristics chart, which displays the mean large-fire size on the Y-axis and mean number of large fires per year on the X-axis. The historical occurrence for the small IPSFC FOA and the larger four-FPU FOA are plotted in **Figure 12**. Because those FOAs have very different sizes, the X-axis is normalized to a one million acre FOA. The curving reference lines indicate lines of equal burn probability (or mean annual area burned per million

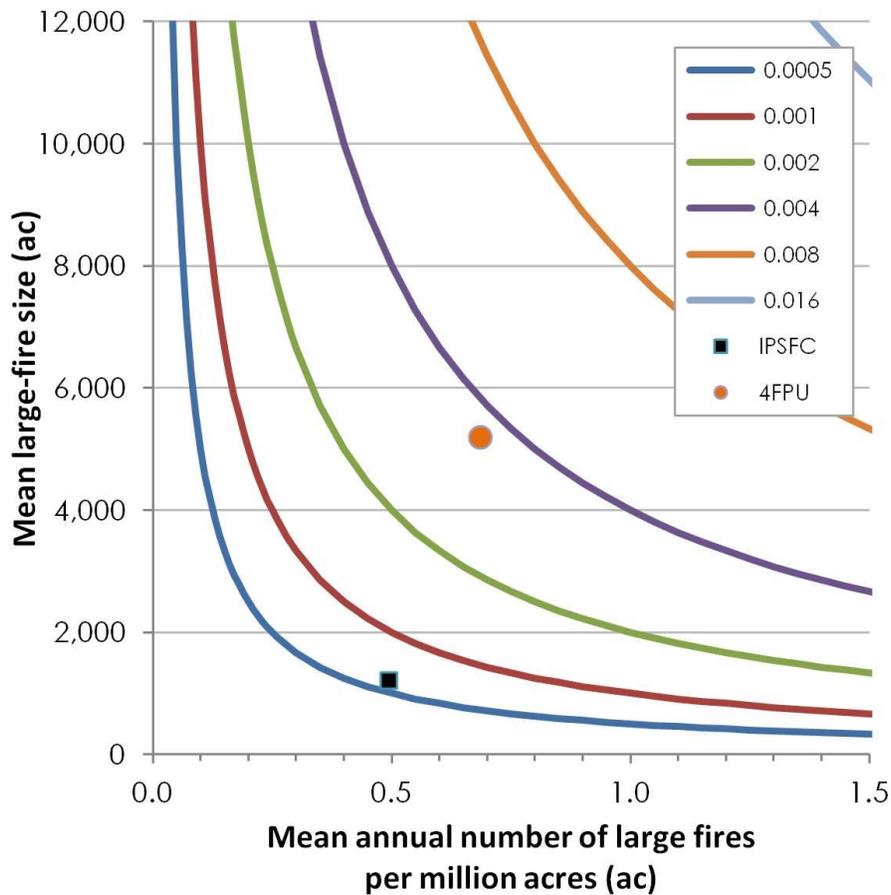


Figure 12. Large-fire occurrence characteristics chart. The Y-axis is the mean large-fire size and the X-axis is the mean annual number of large fires (per million acres of FOA). The curving reference lines represent lines of equal burn probability.

acres). From this chart we can quickly see that the smaller IPSFC FOA has fewer and smaller large fires than the four-FPU FOA as a whole; it therefore also has a much lower mean burn probability.

Although the chart as shown compares two FOAs, it is also useful for comparing the historical results for a FOA with simulated results using FSim. That provides a quick visual cue about what adjustments to inputs are needed to improve agreement between simulated and historical large-fire occurrence.

2.3.7 Ignition density grid

The start locations of wildfires are not necessarily uniform across a FOA. Instead, wildfire start locations may be clustered in certain portions of the FOA due to proximity to ignition sources (for example, human activity centers, or land areas prone to lightning), fuelbed receptivity, climate, and other factors. The locations of fire start locations of any size is an important factor when planning preparedness, prevention and presuppression (initial attack) activities. However, when focusing on the large fires that burn most of the land area, the start locations, though still clustered, may show a different spatial pattern than all fires. The start locations of large fires depend on fuel conditions (some fuel types are prone to rapid spread), remoteness from suppression resources, and other factors such as resistance to control (suppression).

Simulation of large-fire occurrence with FSim is improved by the use of an ignition density grid (IDG). An IDG is a raster data layer that indicates the relative *spatial* likelihood (probability) of large-fire occurrence. (The *temporal* likelihood of a large-fire is handled with the logistic regression model described in section 2.3.9.)

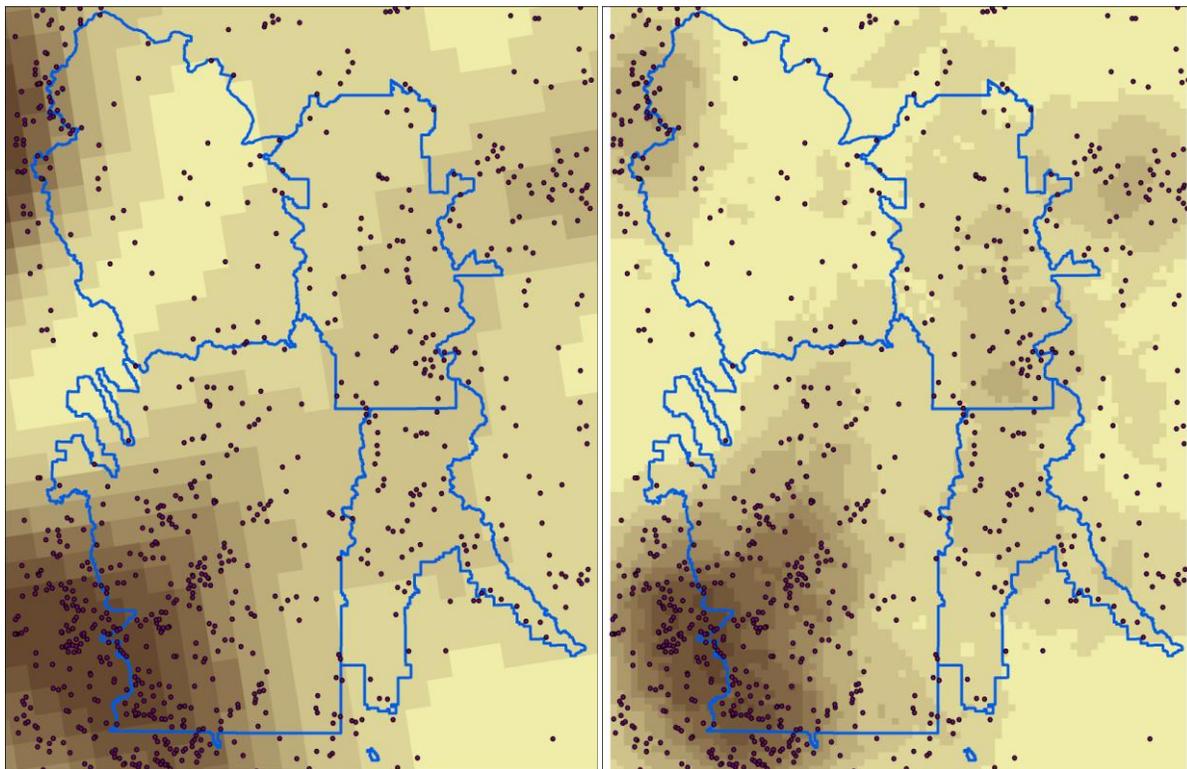


Figure 13. Ignition density grids (IDGs) for the region around the four-FPU FOA. On the left is the 20-km cell size (75 km search radius) IDG developed for use in FPA; on the right is a 5-km cell size (50-km search radius) IDG. Darker colors indicate higher likelihood of large-fire occurrence.

Several methods are available for generating an IDG. For the FPA project, a coarse IDG was developed for the continental US. This raster represents the relative density of large-fire start locations using a 75-km radius around a 20 km grid cell size (**Figure 13**, left panel). The cell with the highest density of starts within the FOA is given a value of 1, and all other cells are scaled as a fraction of that.

In support of a wildfire hazard and risk assessment on the Bridger-Teton National Forest, Sean Parks (of the Aldo Leopold Wilderness Research Institute) built a logistic regression model of large-fire occurrence likelihood based on the factors that affect large-fire occurrence described above—nearness to human activity centers (roads and trails), fuel and vegetation types, etc.

A third method is to generate a custom IDG using methods similar to the FPA IDG, but using a finer grid and smaller search radius. **Figure 13** (right panel) shows the results of running the ArcMap Point Density tool (5 km cell size and 50 km search radius) and then scaling the maximum density to a value of 1.

Because large fires tend to travel long distances from their ignition locations, small differences in IDGs may not have a great effect on the results.

2.3.8 Fire-day distribution

More than one large fire may occur on a single day, especially in a large FOA or one with a lot of wildfire activity. Large-fire occurrence is generally simulated as the likelihood of *at least one* large fire occurring on a given day in relation to a measure of fuel dryness, such as ERC-G (more on this in section 2.3.9 below). A day on which one or more large fires were discovered is called a large-fire day. Accurate simulation of overall large-fire occurrence requires knowledge of the distribution of the number of large fires per large-fire day. This distribution can be determined from the FOD.

For the four-FPU FOA, there were 399 large fires, but only 314 large-fire days. The distribution of the number of large fires per large-fire day is shown in **Figure 14**. Just one large fire occurred on 265 out of 314 (84 percent) of the large-fire days. The mean number of large fires per large-fire day can be calculated by dividing the total number of large fires into the total number of large-fire days. For the four-FPU FOD, that is 399/314, or an average of 1.27 large fires per large-fire day.

2.3.9 Large-fire probability coefficients

In section 2.3.8 above we mentioned that large-fire occurrence is generally simulated as the likelihood of having at least one large fire on a given day. That simulation is most often modeled as a logistic regression based on whether a day in the record was a large-fire day and the ERC-G of each day. The logistic regression can be accomplished with the FireFamilyPlus software or with custom

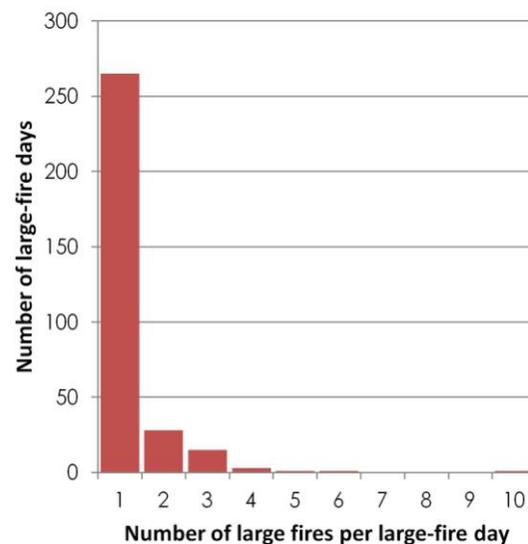


Figure 14. Distribution of the number of large fires per large-fire day for the four-FPU FOD.

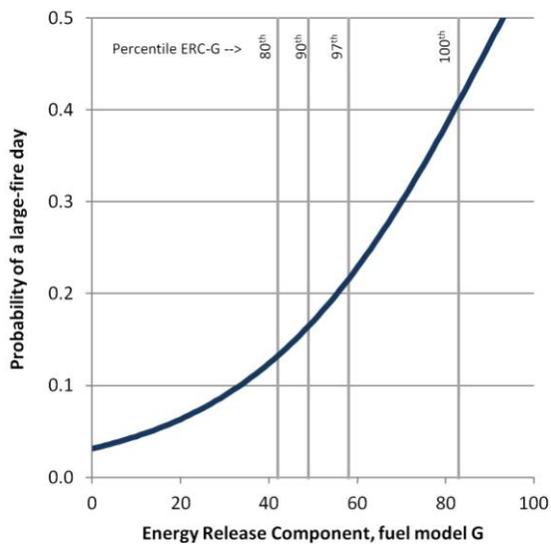


Figure 15. Probability of a large-fire day in relation to the Energy Release Component (ERC) for fuel model G. Reference lines indicate certain year-round percentile ERC values.

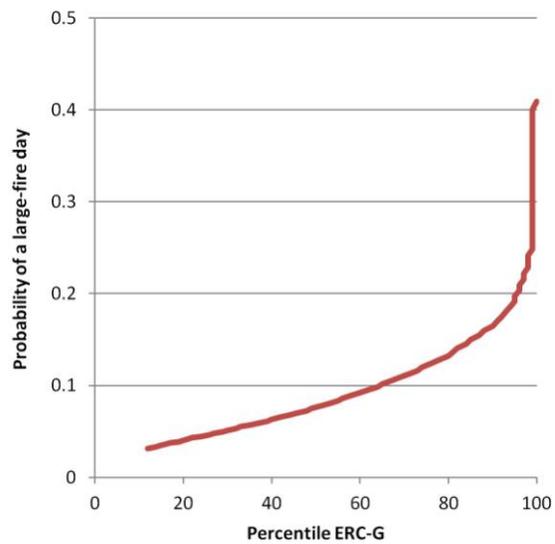


Figure 16. Probability of a large-fire day in relation to percentile ERC-G for the four-FPU FOD and the Island Park, Idaho, RAWS.

scripting in a scripting language such as 'R'.

For the four-FPU FOD, logistic regression coefficients were developed with FireFamilyPlus, and were associated with the ERC-G as measured at the Island Park, Idaho, RAWS station. The coefficients plot out as shown in **Figure 15**. The 80th, 90th, 97th and 100th percentile ERC-G values are referenced on the graph. These results are plotted in terms of the year-round percentile of ERC-G in **Figure 16**. Large-fire growth is commonly associated with ERC-G at or above the 80th percentile, so the probability of a large-fire day ranges approximately 15-20 percent on those days. These probabilities apply to the likelihood of at least one large fire anywhere in the entire 30.8 million acre FOA. The probability of a large-fire in a smaller area would be much lower.

2.4 Summary

In section 2 we described the analysis of contemporary historical wildfire occurrence data. The preparatory steps are to delineate the fire occurrence area (FOA) for the project and to obtain or create a fire occurrence database (FOD) for the delineated FOA.

Once the preparatory steps have been completed, the analysis of historical wildfire occurrence can proceed. That analysis produces 9 results:

1. Lorenz curve
2. Empirical cumulative distribution function
3. Large-fire size threshold
4. Large-fire season chart

5. Seasonal chart of fire occurrence
6. Contemporary large-fire occurrence rates
7. Large-fire ignition density grid
8. Large-fire day distribution
9. Logistic regression coefficients

The first eight of these results can be generated with only the FOD for the FOA. The last result—logistic regression coefficients—also requires a database of daily values of the Energy Release Component (ERC) of the National Fire Danger Rating System (NFDRS) for fuel model G.

These results are useful for planning for future large fires in the FOA, and for parameterizing and calibrating a stochastic large-fire simulator like FSim.

3. Glossary

- **empirical cumulative distribution function (ECDF)**
The empirical cumulative distribution function (empirical CDF) is the CDF generated from an empirical dataset (as opposed to a mathematical model of the data). Like the Lorenz curve, an ECDF provides information about the shape of a distribution of values.
- **fire modeling landscape (LCP)**
A fire modeling landscape (LCP) is a raster (gridded) representation of the fuel, vegetation and topography across a project area. If it is large enough, an LCP extent can serve as a fire occurrence area (FOA).
- **fire occurrence**
Fire occurrence can refer to an individual wildfire event (an instance of a wildfire) or to the analysis of historical wildfire events
- **fire occurrence area (FOA)**
A fire occurrence area is a geographic area for which fire occurrence will be summarized.
- **fire occurrence database (FOD)**
A fire occurrence database is a dataset listing each wildfire event (occurrence) in a specified area. The minimum attributes of a FOD are the start location, start (or discovery) date and final size.
- **large-fire day**
A large-fire day is a day on which one or more large wildfires starts or is discovered.
- **Lorenz curve**
A Lorenz curve illustrates the relationship between the cumulative number of things and their cumulative effects. For wildfire occurrence, the Lorenz curve illustrates the relationship between the number of wildfires (smallest to largest) and their area burned.

4. References

Finney and others 2011

Jolly personal communication

Short 2014

Strauss 1989