

1 **Constructing a database of wildfire evacuations in California 2000-2022**

2 Sam Heft-Neal^{1*}, Lela Hanson², Atsada Israpanich², Talia Buckhouse³, et al tbd

3 ¹Center on Food Security and the Environment, Stanford University, Stanford, CA 94305

4 ² Stanford University, Stanford, CA 94305

5 ³ The Nueva School, San Mateo, CA 94403

6 * Corresponding author contact information:

7 Sam Heft-Neal

8 sheftneal@stanford.edu

9 364 Y2E2

10 Stanford University

11 Stanford, CA 94305

12 **Abstract:** As wildfire activity has accelerated in recent years, evacuations due to fire risk
13 have increased in frequency. However, because evacuations are carried out by a variety of
14 local agencies, there is no centralized data source for evacuation information, limiting re-
15 searchers ability to study evacuations at a broader scale. To address this gap we created a
16 spatio-temporal database of wildfire evacuation orders for all California fires larger than
17 75 acres from 2000-2022. In total, evacuation orders corresponding to 983 fires were identi-
18 fied using systematic searches across inciweb, government agency outlets, media stories, and
19 social media networks (approximately 40% of fires evaluated). For fires with detected evac-
20 uations, daily evacuation timelines were constructed. The compiled information was then
21 used to map spatial extents of each evacuation and their evolution over time. The resulting
22 database includes shapefiles of polygons corresponding to evacuation order boundaries with
23 associated data indicating the fire that caused the evacuation, the evacuation start and end
24 date, the evacuation type (warning, voluntary, or mandatory), the original evacuation or-
25 der text, and links to the source information. This database offers the most comprehensive
26 statewide record of wildfire evacuation order spatial extents in the US. The data can facil-
27 itate research on historical evacuation decisions and responses and in turn inform future
28 evacuation decision making.

29 **Methods**

30 To assemble a spatially explicit database of wildfire evacuations we began with a list of eligible
31 fires, determined whether we could detect evidence of evacuations associated with each fire, and
32 then for those fires with associated evacuations we assembled a collection of textual information
33 before finally mapping the spatial extent of the evacuation orders. In most cases evacuation or-
34 ders for a single fire evolve over time expanding and contracting spatial extents while also tran-
35 sitioning between voluntary, warning, and mandatory evacuation status. For the purpose of this
36 database we aimed to capture these dynamics by creating a new entry into the database (and as-
37 sociated spatial extent) each time a change was detected. An overview of the systematic process
38 used to construct the database is shown in Figure 1 and described in more detail below.

39 **Identifying eligible wildfires.** Our initial sample of fires comes from CAL FIRE’s Fire and Re-
40 source Assessment Program (FRAP). This database of California fires is derived from reports
41 issued by CAL FIRE, the United States Forest Service Region 5, the Bureau of Land Manage-
42 ment, and the National Park Service. The agencies then jointly developed a fire perimeter GIS
43 layer detailing fire perimeters for all reported fires on both public and private lands. Each FRAP
44 observation details key fire information including fire name, incident number, start date, end date,
45 total acreage, and a spatial polygon detailing the fire perimeter.

46 **Assessing whether a fire resulted in evacuation orders.** Once we had assembled a list of el-
47 igible fires we followed a systematic process to check for evidence of associated evacuations.
48 Namely, we first used Google Searches in the form [“<fire name>” “<year of start date>” evac be-
49 fore: <one month after end date> after: <one month before start date>]. If no results were found,
50 the formula was modified at least twice; (1) searching for fire evacuation orders by area instead of
51 the fire name, and (2) removing “evac” from the search term. For fires that started after the year
52 2010, Facebook was also used for investigation, using the formula [<fire name> Fire] and filtering
53 by year. Additionally, for fires that started after the year 2015, X (formerly known as Twitter) was
54 also used for the same purpose, using the formula: [#<fire name>Fire].

55 Because the expansion of social media greatly improved our ability to detect historical evacua-
56 tions, comparisons made over time, and particularly comparisons between time periods before
57 and after social media expanded, using the database should be made with caution.

58 Evacuations were deemed positive after appearing in one source, but negative cases were only
59 marked if there was no information or trace of evacuations after perusing every available source.
60 Campground and trail closures were excluded from these cases because no permanent residents

61 of the area were evacuated.

62 Fires with no evidence of evacuations but whose burned area perimeters overlapped with pop-
63 ulated areas (indicated by a population index greater than 4) were marked as borderline cases.
64 Each fire marked as a negative was then reviewed for a second time by a different person through
65 the same procedure. A total of 983 fires, 36% of the initial list, were found to have associated
66 evacuations.

67 **Compiling evacuation timelines.** We investigated each fire with detected evacuation orders we
68 collected as much publicly available online sources mentioning the associated evacuations as pos-
69 sible. Official sources were prioritized including the United States Incident Information System,
70 government agencies, and Cal Fire. When information was not available from these sources infor-
71 mation was collected first from news articles, and finally, from social media posts. Social media
72 posts from official agencies (e.g., an X post from a Sheriff office's official account) were treated
73 as official sources. 259 fires, or 26% of the list of fires initially marked as potentially having
74 caused evacuations, were found upon further investigation to not have substantial associated evac-
75 uation orders. Many of these false positives were conflated with another fire by the same name or
76 a fire that was occurring simultaneously.

77 Each change in evacuation status on a single day associated with a particular fire was recorded
78 and the source used to identify that change was referenced. Changes in evacuation status include
79 both changes in the size of the area under evacuation or changes in the evacuation level (volun-
80 tary/warning/mandatory/lifted). Multiple changes in a single day were combined in chronological
81 order where the issuance of evacuations was prioritized over evacuation lifts.

82 In some cases when evacuation orders were detected, we could not find evidence of the evacua-
83 tion order being lifted. In these cases we left the variable indicating the end date of the evacua-
84 tion missing.

85 **Mapping evacuation boundaries.** Finally, we mapped the spatial extent of each evacuation or-
86 der status. Spatial information regarding evacuations comes in numerous forms which we divided
87 into four categories:

- 88 1. Maps or zones - when the evacuation order includes a map or references specific zones as-
89 sociated with defined geographies.
- 90 2. Neighborhoods - when the evacuation order references neighborhood names.
- 91 3. Geometric areas defined by landmarks - when the evacuation order infers a geography by

92 listing land marks as boundary points.

93 4. Miscellaneous - when the evacuation order includes other information not falling into cate-
94 gories 1-3.

95 QGIS, a free and open-source platform geographic information system application, was used to
96 map the evacuation information compiled in the timelines. OpenStreetMap and Google Earth
97 were used to identify geographic landmarks, while the fire perimeter shapefiles from the Fire and
98 Resource Assessment Program were used as references. Each change in evacuation status was
99 assigned a unique evacuation identification for that particular fire. Each polygon had eight as-
100 sociated fields: (1) id, (2) type, (3) start date, (4) end date, (5) last known, (6) evac id, (7) notes,
101 and (8) desc. We recorded the associated fire's identification number under "id", and assigned
102 an identification number to each evacuation change for a particular fire under "evac id"; "last
103 known" was used exclusively for cases where no information was found about the end date of
104 the evacuation; "notes" was used to elaborate on ambiguity and indicate significant evacuation
105 failures; "desc" was used to record the original textual description of the evacuation area.

Figure 1: Diagram of process identifying which fires had evacuations.



