CSIRO PUBLISHING

International Journal of Wildland Fire
http://dx.doi.org/10.1071/WF16050

Research Note

Historical reconstructions of California wildfires vary by data source

Alexandra D. Syphard[A,D] *and Jon E. Keeley*[B,C]

[A]Conservation Biology Institute, 10423 Sierra Vista Avenue, La Mesa, CA 91941, USA.
[B]US Geological Survey, Western Ecological Research Center, Sequoia–Kings Canyon Field Station, 47050 Generals Highway, Three Rivers, CA 93271, USA.
[C]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA.
[D]Corresponding author. Email: asyphard@consbio.org

Abstract. Historical data are essential for understanding how fire activity responds to different drivers. It is important that the source of data is commensurate with the spatial and temporal scale of the question addressed, but fire history databases are derived from different sources with different restrictions. In California, a frequently used fire history dataset is the State of California Fire and Resource Assessment Program (FRAP) fire history database, which circumscribes fire perimeters at a relatively fine scale. It includes large fires on both state and federal lands but only covers fires that were mapped or had other spatially explicit data. A different database is the state and federal governments' annual reports of all fires. They are more complete than the FRAP database but are only spatially explicit to the level of county (California Department of Forestry and Fire Protection – Cal Fire) or forest (United States Forest Service – USFS). We found substantial differences between the FRAP database and the annual summaries, with the largest and most consistent discrepancy being in fire frequency. The FRAP database missed the majority of fires and is thus a poor indicator of fire frequency or indicators of ignition sources. The FRAP database is also deficient in area burned, especially before 1950. Even in contemporary records, the huge number of smaller fires not included in the FRAP database account for substantial cumulative differences in area burned. Wildfires in California account for nearly half of the western United States fire suppression budget. Therefore, the conclusions about data discrepancies and the implications for fire research are of broad importance.

Additional keywords: area burned, data uncertainty, fire frequency, fire perimeters, historical trends.

Received 10 November 2015, accepted 7 August 2016, published online 19 October 2016

Introduction

Wildfire activity has fluctuated dramatically during the last century and fire regimes have been altered in different directions and magnitudes, depending upon geographic region (Safford and Van de Water 2014). Understanding the drivers and nature of these changes, thereby projecting the future trajectory of wildfire activity, has critical implications for the ecological structure and function of fire-prone ecosystems (Pausas and Keeley 2009), long-term carbon stocks (Bachelet *et al.* 2015), and human welfare and safety (Syphard *et al.* 2012). It is therefore essential to attain and analyse historical patterns of wildfire activity to facilitate research on drivers of change, conduct risk assessments and prioritise different management operations (e.g. Hardy and Hardy 2007; Thompson and Calkin 2011; Miller *et al.* 2012). In other words, to understand and anticipate future conditions, it is necessary to study patterns and trends in the past (Safford *et al.* 2012), and this depends on accurate and complete records of past fire events.

Records of wildfire activity historically have been maintained in written or mapped formats by government agencies, each with different formatting requirements and standards. The primary concern has been the completeness of the record, but that is only one of many potential uncertainties (Regan *et al.* 2002). Given such disparate sources and formats of fire data, it has been challenging to assemble complete and reliable historical datasets for areas spanning multiple land ownerships over long time periods (Short 2014).

California is noteworthy for its extensive written fire history for state and federal lands (Keeley and Syphard 2015). These data are available as annual summaries of fire frequency and area burned for United States Forest Service (USFS) forests (1910–present) and for individual counties protected by the state agency California Department of Forestry and Fire Protection (Cal Fire) (1919–present). These data include all fires and their size that were recorded each year by county (Cal Fire) or forest (USFS). Another database is derived from fire perimeter data reported annually for moderate–large fires by state and federal agencies and is known as the State of California Fire and Resource Assessment Program (FRAP) fire history database. One advantage of this geographic information system (GIS)-format database is that it delineates the perimeters of individual fires but it is restricted to larger fires, where information on their spatial location is available. This FRAP database has been used extensively in the study of historical patterns of burning in the state (e.g. Keeley *et al.* 1999; Oneal *et al.* 2006; Moritz *et al.* 2009; Schwartz *et al.* 2015) or how historical fires have affected vegetation patterns (e.g. Franklin *et al.* 2004; Talluto and Suding 2008; Meng *et al.* 2014) to examine both drivers in annual variability and ecosystem impacts of fire activity.

Given the importance of historical fire data for understanding trends and drivers of fire activity, our objective was to better understand the characteristics and differences between these two historical sources of fire data in the state of California through a historical comparison of fire activity (1919–2013). We also compared these two data sources with a more contemporary dataset (1992–2013) known as the national interagency Fire Program Analysis, Fire-Occurrence Database (FPA FOD) (Short 2014).

We asked:

1) Are there differences in total number of fires and area burned over time among the datasets?
2) Do data discrepancies vary by climatic division?
3) How synchronous is variation in modern records of fire activity among databases?
4) What attributes does each of these datasets possess that reflect their most appropriate application?

Methods

Given that fire regimes and their drivers are spatially variable (Keeley and Syphard 2015), we stratified our comparative analysis among the most fire-prone climatically-homogeneous divisions in the state (Fig. 1) using the National Oceanographic and Atmospheric Administration's (NOAA's) boundaries (http://www.esrl.noaa.gov/psd/data/usclimdivs/data/map.html, accessed 27 August 2015). Of the seven divisions spanning the state, we compiled data for the North Coast Drainage (North Coast), Sacramento Drainage (North Interior), San Joaquin Drainage (Sierra Nevada), Central Coast Drainage (Central Coast) and the South Coast Drainage (South Coast). Due to overlapping boundaries of US Department of Agriculture (USDA) Forest Service lands across the South Coast and Central Coast drainages, and because these divisions are characterised by similar climate and vegetation patterns, we grouped these two divisions together and refer to them as 'South-Central Coast'.

FRAP database

These data are derived from individual fire events, and fire perimeters are digitised from fire maps available from state and federal agencies and some local agencies, each with associated identifiers such as fire name, date and size. The fire size measurement includes all area within the perimeter, although large fires often contain unburned patches. Initially, the database covered 1950–2001 and included wildland fires larger than 4 ha

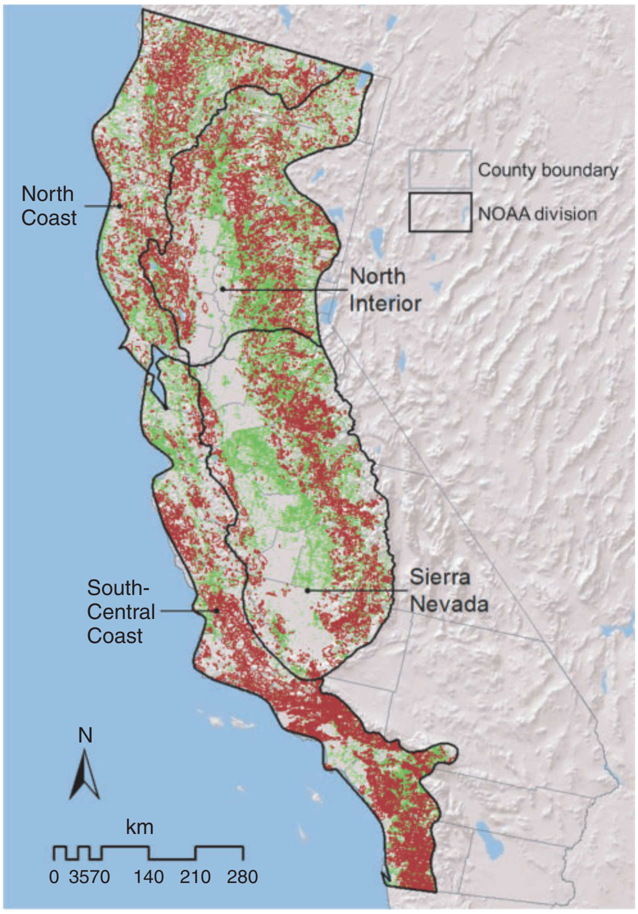

Fig. 1. The study area showing National Oceanographic and Atmospheric Administration (NOAA) climate divisions in California, USA, and location of fires for the two spatial datasets, the State of California Fire and Resource Assessment Program (FRAP) fire history database (FRAP, red perimeters) and the National interagency Fire Program Analysis, Fire-Occurrence Database (FPA FOD, green points).

on USDA Forest Service land and fires larger than 121 ha on lands under the protection of Cal Fire but many exceptions are apparent. Beginning in 1989, Cal Fire began annually updating the database by adding all fire information back to 1878, identifying and removing duplicate fires, and adding fires from the previous season. Nevertheless, many fires may be missing either because they were not mapped or were missing documentation, a limitation that is more common as one goes back in time.

The FRAP fire perimeter database, updated to 2013, was downloaded from (http://frap.fire.ca.gov/data/frapgisdata-sw-fireperimeters_download.php, accessed 28 August 2014). We only summarised data for wildfires and excluded prescribed burns. To match the earliest date of complete fire records available through the written records for both Cal Fire and USFS, we selected all fires in the FRAP database that occurred from 1919 to 2013 and clipped them to the boundaries of the climate divisions. We limited our analysis to fires or portions of fires that were within those areas to ensure complete geographical overlap with the data reported in the written records. In the case that the administrative boundary crossed the climate division boundary along the eastern side of the state (e.g. Fig. 1), we carved out the extent of fire perimeters within those areas to ensure complete geographical overlap with the data reported in the written records.

Written data

We assembled the written records of annual wildfire activity from the two sources with the longest records in the state: 1) USFS fire data that covered all national forests for the years 1910–2013 (although we used only the years 1919–2013) and 2) Cal Fire data that covered direct protection areas (DPA), which are mostly state responsibility lands, for the years 1919–2013 (see Keeley and Syphard 2015). These written data differ from the FRAP database in that they include all wildfires regardless of size but they are only spatially explicit to the level of forest (USFS) or county (Cal Fire).

Most of fires that comprise the Cal Fire and USFS fire databases are only represented in one or the other database. Very large fires are often included in both, theoretically agencies only report the acreage burned within their jurisdiction. Nonetheless, the potential exists for double reporting of area burned, so to estimate this potential source of error we did the following more detailed investigation. For years 2003–2008, we examined individual fires that exceeded 4000 ha to determine which had been reported by both agencies and to what extent the combined Cal Fire and USFS area burned data exceeded the reported total fire size. These years represented some of the largest fires in southern California (2003, 2007) and northern California (2008) and thus would be more prone to double reporting than would be typical through the entire record.

Contemporary FPA FOD database

We compared these historical datasets with FPA FOD database for the 20-year period of overlap. These include all available fire data from local, state and federal sources. We downloaded these data as a GIS layer that includes a point location within a 260 ha (1 square mile) grid for every fire, with fire size provided as a digital attribute (http://www.fs.usda.gov/rds/archive/Product/RDS-2013-0009.3/, accessed 27 August 2015). We clipped all GIS points from the database to the four NOAA division boundaries and subsequently summed the total number of fires and area burned for each year, regardless of agency source.

Analysis

Because we were only interested in the completeness of each data source with respect to fire frequency and area burned, we performed simple summaries of each dataset by year and by climate division for comparison. The spatial nature of each dataset is so fundamentally different that a spatial comparison finer than the unit of NOAA climate division was not possible.

We also summarised the annual number of small fires, i.e. ≤4.047 ha, within each climate division for the two historical datasets and calculated their proportion of the total number of fires for which size class data were available. These summaries allowed us to visualise the proportion of small fires that may have been missed on account of size limitations in the FRAP data. For the FRAP data, we identified small fires by calculating the area of each fire perimeter. For the written records, we summarised fires that were recorded to be in size classes A (<0.101 ha/quarter of an acre) or B (>0.101 ha/quarter of an acre but <4.047 ha/10 acres), which consistently represented all fires ≤4.047 ha.

Results

Fire frequency

The number of fires in the written record greatly exceeded, by at least one order of magnitude for most divisions, the number reported by the FRAP database in all four climate divisions (Fig. 2*a*–*d*). The discrepancy in the South-Central Coast Division was generally smaller than in the other divisions (Fig. 2*d*). The largest discrepancies occurred in the latter half of the 20th century, but the two datasets started to converge in the last 10 years in all regions. Cumulatively, the discrepancies in fire frequency resulted in more than one order of magnitude difference in all divisions except the South-Central Coast, where the written record nevertheless contained nearly 6 times the number of fires than the FRAP record (Fig. 3*a*).

The discrepancy in the proportion of small fires was generally highest in the latter part of the 20th century, with the written record containing a larger proportion of small fires than FRAP (Fig. 4). Nevertheless, the discrepancy in small fires was generally not as large as that in the total number of fires (Fig. 2) and the convergence in the last 10 years was less evident.

The more contemporary FPA FOD database reported fewer fires than the written record and more fires than the FRAP record until ~2007, after which fire frequency either equalled (North Coast and North Interior) or slightly exceeded (Sierra Nevada and South-Central Coast) the frequency reported by the other datasets (Fig. 2). Cumulatively, the FPA FOD data reported fewer fires than the written record but more fires than FRAP during this period (Fig. 3*b*).

Area burned

In the evaluation of potential error due to double reporting of area burned, we found that the 2003 Cedar Fire, the largest in our database, had a reported area burned by Cal Fire of 59 169 ha and by USFS of 69 321 ha, which was 15 150 ha or 13% higher than the total fire size of 113 340 ha. However, this was an anomaly for fires during the years 2003–2008. Over that time period, there were 89 fires larger than 4000 ha, with only 13 reported by both agencies. These 89 fires comprised a total area burned of 1 339 414 ha, and for these fires, the combined Cal Fire and USFS area burned was 3.5% higher. Considering that there were many more large fires during this time period than throughout the rest of the record, it is likely that overall the potential error in the written records is less than 3.5%. These years represented some of the largest fires in southern California (2003, 2007) and northern California (2008) and thus would be more prone to double reporting than would be typical through the entire record. Both state and federal agencies are not likely to respond to small fires and indeed our statistics overwhelmingly support that conclusion – that is, most fires of more than 4000 ha are attacked by just one or the other agency. However, from 2003 to 2008 many large fires occurred and both agencies were likely to be involved.

Most divisions showed substantially more area burned for written records than for FRAP, especially in the first half

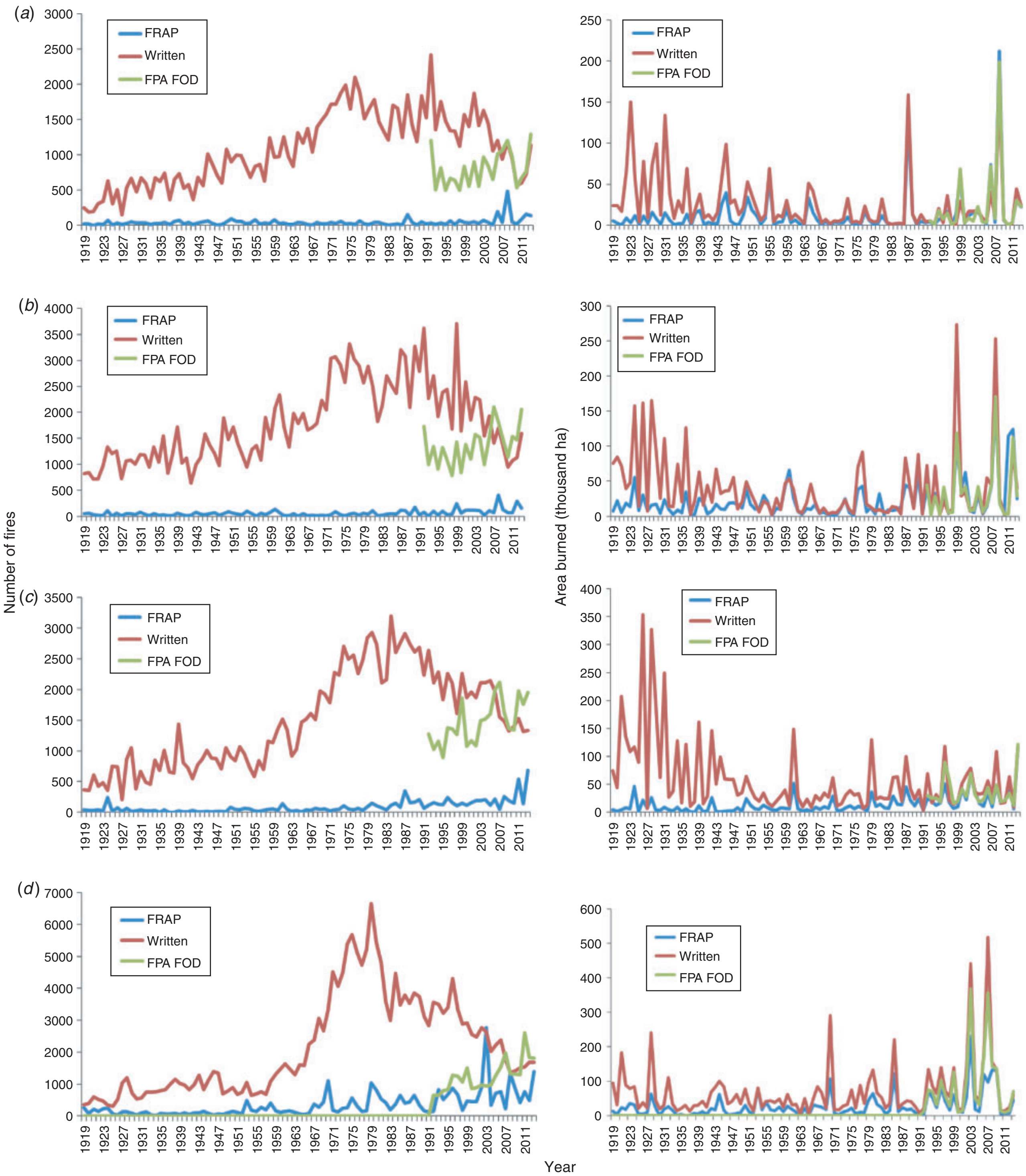

Fig. 2. Number of fires and area burned over time for three datasets in the (*a*) North Coast, (*b*) North Interior, (*c*) Sierra Nevada and (*d*) South-Central Coast divisions of California, USA. FRAP is California's Fire and Resource Assessment Program; FPA FOD is the national interagency Fire Program Analysis, Fire-Occurrence Database.

of the 20th century (Fig. 2). This was most pronounced in the northern half of the state (Fig. 2*a–c*). After ~1950, all divisions showed diminished discrepancy between FRAP and the written record.

In the last two decades, the area burned reported in the FPA FOD database generally matched that reported in the two historical databases. However, the written records tend to exceed the area reported in either FPA FOD or FRAP databases,

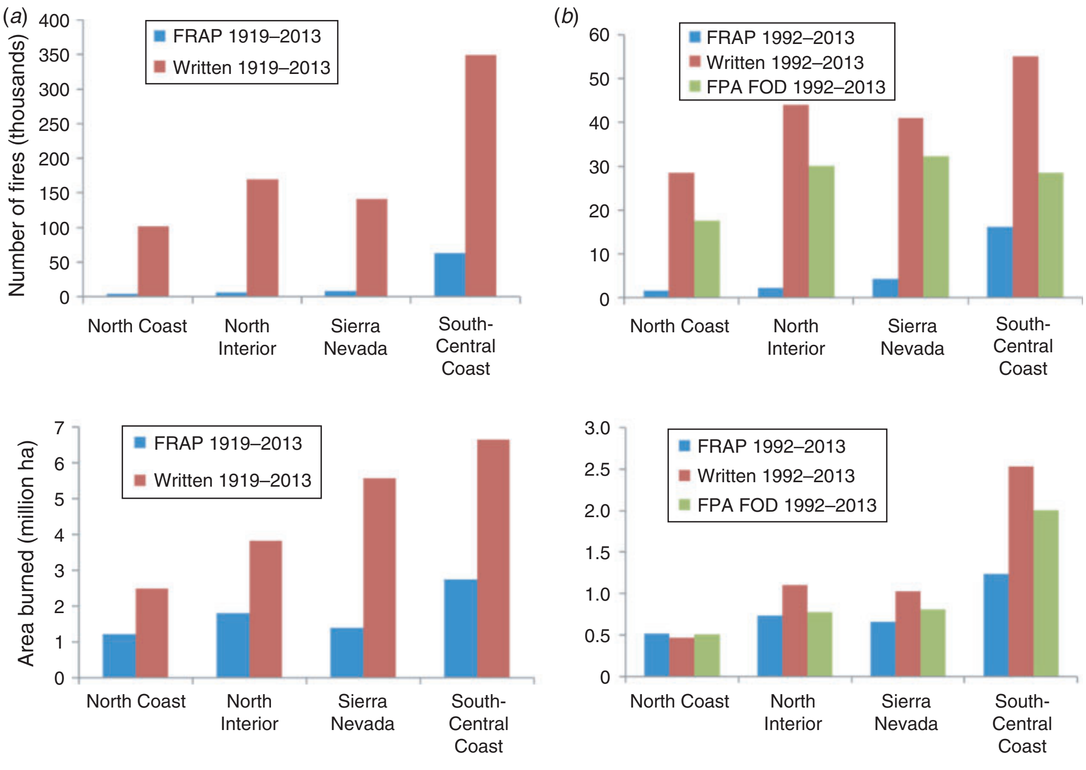

Fig. 3. Cumulative number of fires and area burned (*a*) from 1919–2013 as reported in the FRAP and written records and (*b*) from 1992–2013 as reported in the FRAP, written, and FPA FOD records.

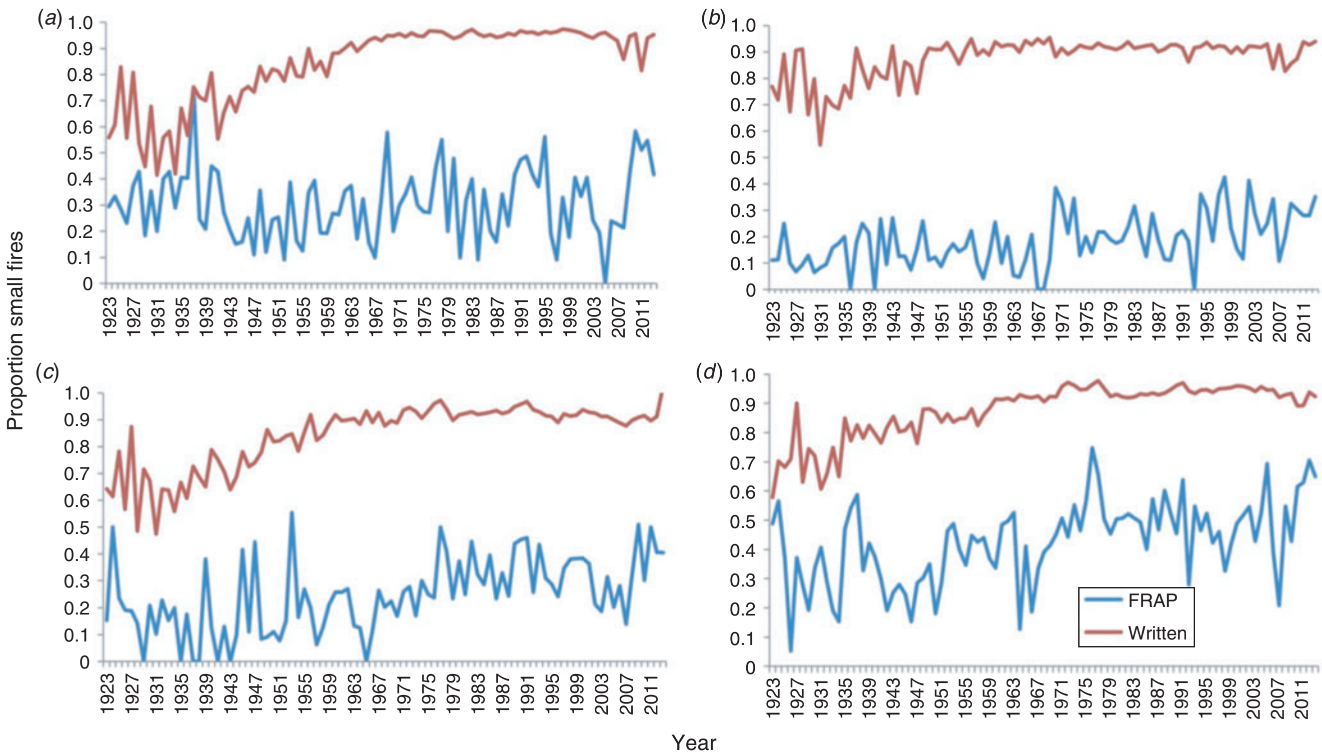

Fig. 4. Proportion of small fires ≤4.047 ha over time in the FRAP and written records for the climate divisions (*a*) North Coast, (*b*) North Interior, (*c*) Sierra Nevada and (*d*) South-Central Coast.

particularly in the North Interior and South-Central Coast drainages, but not in the North Coast (Fig. 2*a*, *b*, *d*). This was the case even in the years in which the FPA FOD data contained fewer fires than the written records (~1992–2007). Nevertheless, across all datasets, peak fire years corresponded nearly identically in all climate divisions.

The cumulative differences in area burned over time were substantial (Fig. 3*a*), with approximately 2–3 times as much area burned reported in the written record than summarised across the FRAP data. These cumulative differences have persisted in the contemporary record (Fig. 3*b*), except for the North Coast Division, where the reported area burned is nearly the same for the three databases. Compared with FRAP, the FPA FOD dataset reported slightly larger cumulative area burned, except for the South-Central Coast where the difference was fairly substantial (Fig. 3*b*). Compared with the written record, the FPA FOD generally reported less area burned, except for the North Coast.

Discussion

The substantial differences in the California historical databases clearly reveal different strengths and weaknesses. The FRAP database is indispensable for tracking the spatial pattern of past fires, as it is the only one presenting fire perimeters. The primary advantage of the written fire record is that it is substantially more complete throughout much of the historical period, particularly in terms of fire frequency. In divisions like the Sierra Nevada, with substantial National Park Service land, this distinction might be even larger because fires on these National Park Service lands were not included in the written database analysed here, but were included in the FRAP record.

One limitation of FRAP is that, for a fire to be included, it requires spatially explicit information to generate fire perimeters, and in the earlier historical record, these were less readily available and undoubtedly less precise than modern fire maps. This is why the FRAP database was initially designed to cover the second half of the 20th century. Another reason for the lower area burned in the FRAP data is the exclusion of small fires. Whereas small fires make up most of total fires in the written record, only about a third to half of those have been included in FRAP (Fig. 4) and the area of these small fires, combined with the area of fires that have not been mapped, becomes substantial over time (Fig. 3).

Appropriate uses of these databases

The written records are most useful for analyses that require the most complete record of 20th century patterns of burning. For example, had the recent study by Keeley and Syphard (2015), relating historical fire-climate patterns for the last century, been based on the FRAP database, it would have grossly underestimated fire activity in the first half of the 20th century and missed important fire-climate relationships. The more complete written records would also be the most appropriate dataset for evaluating how ignition sources have changed over time because, until recently, the FRAP and the FPA FOD databases greatly underestimated ignitions (Fig. 2). However, if one were interested only in ignition sources for large fire events, both FRAP and the FPA FOD would be useful for data since 1950 (FRAP) or more recently (FPA FOD). Recognising this potential for underreporting small fires in the FRAP database, Miller *et al.* (2012) restricted an historical analysis of fire activity to fires >40 ha from 1910–2008, but even some large fires may have been missed.

Because the FRAP database has been used by numerous investigators, it is worth evaluating potential shortcomings that may not have been understood by those investigators. One of the first scientific uses of this database was the 1999 report on fire suppression impacts in southern California (Keeley *et al.* 1999). By using the FRAP database, that study greatly underestimated the extent of burning during the first half of the 20th century and underestimated the marked drop in fire size in recent decades. If the written records had been used, the conclusions about minimal fire suppression impacts in southern California would have been even stronger.

In their study of fire frequency on alien plant distribution in Orange County, California, Talluto and Suding (2008), using the FRAP database, attributed some increases in alien plants to nitrogen deposition as opposed to fire history. While this link may be valid, the limitations of the FRAP database, particularly in failing to record most of small fires, makes it likely that fire disturbance is of greater importance than they concluded. In this county, more than 95% of the fires reported in the written record are not recorded in the FRAP database, so the potential exists for vast portions of this county having experienced repeated fires not being recorded in the FRAP database. Similar problems may exist in the quantification of fire intervals using the FRAP data in Moritz *et al.* (2009) and Meng *et al.* (2014).

Krawchuk and Moritz (2012) relied on the FRAP database for their fire-climate analysis in California. Although much of the work focused on recent decades where the FRAP database is comparable to other databases in area burned (Fig. 2), they made comparisons with other time periods beginning in 1878, 1911 and 1941, and they used the FRAP database to investigate patterns of ignitions. While they fully acknowledge limitations of these earlier datasets, the findings here suggest that their conclusions may need re-evaluation with the more complete written data source (Fig. 2).

The FRAP database has also been used to explain successional changes (Franklin *et al.* 2004). In this sort of approach, scale is very important since, if the analysis is done over a large enough area, there is the potential for small errors to even out. Nevertheless, for any given site, the FRAP stand age should be questioned. For example, of 180 sites measuring 0.1 ha, Keeley *et al.* (2008) found that the actual age of the stand based on stem ring counts was younger than the FRAP age 53% of the time. In short, FRAP fire perimeter maps should be considered an unreliable indicator of the age of a particular site.

A recent analysis using the FRAP database contends that, since 1911, the distribution of fires in the Sierra Nevada of California have risen in elevation, a potential sign of climate change impacts (Schwartz *et al.* 2015). Since this study shows a clearly demonstrable change in the reliability of the FRAP data beginning in the mid-20th century (Fig. 4), these data are clearly inappropriate for demonstrating historical trends back to the early 20th century. Nevertheless, since fire-perimeter maps are required for this type of study, the FRAP database is the only viable option. The Schwartz *et al.* (2015) paper did a breakdown of change by era, suggesting that the change in elevation has continued to the present. Thus, the overall conclusions of that paper are likely not affected by dependence on early FRAP records.

Conclusions

The evaluation considered here broadly applies to historical reconstructions. In general, written records are the most complete if one is concerned with long-term fire trends. The major advantage of the FRAP database is that it provides spatially explicit fire perimeters that are required for many purposes. Since 1950, the FRAP database does a fine job capturing the largest fires and the FAA FOD database also does this for the last few decades. However, FRAP is not designed to capture small fires, which even today comprise a substantial proportion of all fires. Thus, on landscapes where small fires are abundant, FRAP may significantly underestimate area burned. It is important that users closely evaluate the limitations of these datasets and their particular data needs before making a selection.

References

Bachelet D, Ferschweiler K, Sheehan TJ, Sleeter BM, Zhu Z (2015) Projected carbon stocks in the conterminous US with land use and variable fire regimes. *Global Change Biology* **21**, 4548–4560. doi:10.1111/GCB.13048

Franklin J, Coulter C, Rey SJ (2004) Change over 70 years in a southern California chaparral community related to fire history. *Journal of Vegetation Science* **15**, 701–710. doi:10.1111/J.1654-1103.2004.TB02312.X

Hardy CC, Hardy CE (2007) Fire danger rating in the United States of America: an evolution since 1916*. *International Journal of Wildland Fire* **16**, 217–231. doi:10.1071/WF06076

Keeley JE, Syphard AD (2015) Different fire-climate relationships on forested and non-forested landscapes in the Sierra Nevada region. *International Journal of Wildland Fire* **24**, 27–36. doi:10.1071/WF14102

Keeley JE, Fotheringham CJ, Morais M (1999) Reexamining fire suppression impacts on brushland fire regimes. *Science* **284**, 1829–1832. doi:10.1126/SCIENCE.284.5421.1829

Keeley JE, Brennan T, Pfaff AH (2008) Fire severity and ecosystem responses following crown fires in California shrublands. *Ecological Applications* **18**, 1530–1546. doi:10.1890/07-0836.1

Krawchuk M, Moritz M (2012) Fire and climate change in California. California Energy Commission Climate Change Center, Public Interest Energy Research Program White Paper. (Sacramento, CA) Available at http://www.energy.ca.gov/2012publications/CEC-500-2012-026/CEC-500-2012-026.pdf [Verified 9 August 2016]

Meng R, Dennison PE, D'Antonio CM, Moritz MA (2014) Remote sensing analysis of vegetation recovery following short-interval fires in Southern California shrublands. *PLoS One* **9**, e110637. doi:10.1371/JOURNAL.PONE.0110637

Miller JD, Skinner CN, Safford HD, Knapp EE, Ramirez CM (2012) Trends and causes of severity, size, and number of fires in northwestern California, USA. *Ecological Applications* **22**, 184–203. doi:10.1890/10-2108.1

Moritz MA, Moody TJ, Miles LJ, Smith MM, de Valpine P (2009) The fire frequency analysis branch of the pyrostatistics tree: sampling decisions and censoring in fire interval data. *Environmental and Ecological Statistics* **16**, 271–289. doi:10.1007/S10651-007-0088-Y

Oneal C, Stuart J, Steinberg S, Fox LI (2006) Geographic analysis of natural fire rotation in the California redwood forest during the suppression era. *Fire Ecology* **2**, 73–99. doi:10.4996/FIREECOLOGY.0201073

Pausas JG, Keeley JE (2009) A burning story: the role of fire in the history of life. *Bioscience* **59**, 593–601. doi:10.1525/BIO.2009.59.7.10

Regan HM, Colyvan M, Burgman MA (2002) A taxonomy and treatment of uncertrainty for ecology and conservation biology. *Ecological Applications* **12**, 618–628. doi:10.1890/1051-0761(2002)012[0618:ATATOU]2.0.CO;2

Safford HD, Van de Water KM (2014) Using fire return interval departure (FRID) analysis to map spatial and temporal changes in fire frequency on national forest lands in California. USDA Forest Service, Pacific Southwest Research Station, Research Paper PSW-RP-266. (Albany, CA)

Safford HD, Hayward GD, Heller NE, Wiens JA (2012) Historical ecology, climate change, and resource management: can the past still inform the future? In 'Historical Environmental Variation in Conservation and Natural Resource Management' (Eds JA Wiens, GD Hayward, HD Safford, CM Giffen) pp. 46–62. (Wiley-Blackwell, Oxford, UK)

Schwartz MW, Butt N, Dolanc CR, Holguin A, Moritz MA, North MP, Safford HD, Stephenson NL, Thorne JH, van Mantgem PJ (2015) Increasing elevation of fire in the Sierra Nevada and implications for forest change. *Ecosphere* **6**, art121. doi:10.1890/ES15-00003.1

Short K (2014) A spatial database of wildfires in the United States, 1992–2011. *Earth System Science Data* **6**, 1–27. doi:10.5194/ESSD-6-1-2014

Syphard AD, Keeley JE, Massada AB, Brennan TJ, Radeloff VC (2012) Housing arrangement and location determine the likelihood of housing loss due to wildfire. *PLoS One* **7**, e33954. doi:10.1371/JOURNAL.PONE.0033954

Talluto M, Suding K (2008) Historical change in coastal sage scrub in southern California, USA in relation to fire frequency and air pollution. *Landscape Ecology* **23**, 803–815. doi:10.1007/S10980-008-9238-3

Thompson MP, Calkin DE (2011) Uncertainty and risk in wildland fire management: a review. *Journal of Environmental Management* **92**, 1895–1909. doi:10.1016/J.JENVMAN.2011.03.015