Review

Sources and implications of bias and uncertainty in a century of US wildfire activity data

Karen C. Short

USDA Forest Service, Rocky Mountain Research Station, Forestry Sciences Laboratory,
800 East Beckwith Avenue, Missoula, MT 59801, USA. Email: kcshort@fs.fed.us

Abstract. Analyses to identify and relate trends in wildfire activity to factors such as climate, population, land use or land cover and wildland fire policy are increasingly popular in the United States. There is a wealth of US wildfire activity data available for such analyses, but users must be aware of inherent reporting biases, inconsistencies and uncertainty in the data in order to maximise the integrity and utility of their work. Data for analysis are generally acquired from archival summary reports of the federal or interagency fire organisations; incident-level wildfire reporting systems of the federal, state and local fire services; and, increasingly, remote-sensing programs. This paper provides an overview of each of these sources and the major reporting biases, inconsistencies and uncertainty within them. Use of national fire reporting systems by state and local fire organisations has been rising in recent decades, providing an improved set of incident-level (documentary) data for all-lands analyses of wildfire activity. A recent effort to compile geospatial documentary fire records for the USA for 1992–2013 has been completed. The resulting dataset has been evaluated for completeness using archival summary reports and includes a linkage to a widely used, remotely sensed wildfire perimeter dataset.

Introduction

The statistical analysis of wildfire activity is a critical component of national wildfire planning, operations, and research in the United States. Wildfire activity data have been collected in the USA for over a century. Yet, to this day, no single, unified system of wildfire record-keeping exists. Data for analysis are generally harvested from archival summary reports of the federal or interagency fire organisations; incident-level wildfire reporting systems of the federal, state and local fire services; and, increasingly, remote-sensing programs. It is typical for research into wildfire activity patterns for all or part of the last century to require data from several of these sources and perhaps others. That work is complicated by the disunity of the various datasets and potentially compromised by inherent reporting biases, inconsistencies and errors or uncertainty in the data, as described here. The extent and quality of data available from any one source may also change considerably over time, and the issues summarised in the present review are important to recognise even for analyses restricted to a single source of wildfire activity data.

Data sources

Data for analyses of variables like wildfire numbers and area burned in the US are available in various forms from the early 20th century to the present. Fig. 1 shows the general temporal coverage of key data sources grouped into three categories: (1) archival summary reports, (2) documentary fire records, and (3) remotely sensed data.

The longest time period is covered by the archival summary reports, ~1916 to 1997, from the US Forest Service (USFS). These reports include annual estimates of wildfire activity on all US lands (forested and non-forested) that qualified for organised wildfire protection in any given year, as per the statutory requirements of the 1911 Weeks Law and subsequent legislation. Figures were reported for each US state and included estimates of wildfire numbers and area burned by land ownership, fire cause and fire size class. The National Interagency Coordination Center (NICC) began issuing an independently sourced set of similar reports in 1983. The basic fire occurrence information in those annual summaries comes from the NICC's Situation Reporting module, which gets daily activity reports from wildland fire dispatch offices during the fire season and weekly reports otherwise. Together, the USFS and NICC archival summary reports provide over a century of wildfire-activity data (Fig. 1), and are commonly used to characterise trends in variables like wildfire area burned (WFAB) over lengthy periods at the national level or by interstate region (e.g. Houghton *et al.* 2000; Mouillot and Field 2005; Collins *et al.* 2006; Littel *et al.* 2009).

Documentary fire records, which are created using standard forms issued by the various federal, state and local fire organisations, are warehoused and readily available in electronic format for several recent decades (Fig. 1). These fire reports are intended to capture the basic facts about each individual wildfire occurrence – including point of origin, date, cause and final fire size – as well as additional information about responses and

impacts associated with the incident. The USFS, bureaus of the US Department of Interior (USDI), state fire authorities and local fire departments use different forms and reporting systems to capture these incident-level data, and their electronic archives span different time periods (approximated in Fig. 1). Some states, like California and Alaska, have digital archives (including locally mapped perimeters of large fires) that extend over longer historical periods than Fig. 1 indicates; but their extensive electronic archives are exceptional cases. The fire–climate analyses of Westerling *et al.* (2003, 2006) are examples of work based on documentary wildfire records.

For recent decades, remotely sensed fire data are also available, including satellite-derived perimeters, or burn scars, of large (i.e. >405 ha) fires dating back to 1984, from the Landsat-based Monitoring Trends in Burn Severity (MTBS) project (Eidenshink *et al.* 2007). The MTBS data have increasingly been used to identify trends in and drivers of large-fire activity in the US (e.g. Riley *et al.* 2013; Dennison *et al.* 2014; Lannom *et al.* 2014). The remote-sensing group (Fig. 1) also includes satellite fire detection data like the active fire (hotspot) and burned area products that date back to ~2000 from use of the Moderate Resolution Imaging Spectroradiometer (MODIS) on board the Aqua and Terra satellites (Giglio *et al.* 2003). Although other satellite sensors have advanced fire detection capabilities, including the Geostationary Operational Environmental Satellite (GOES) system and the Visible Infrared Imaging Radiometer Suite (VIIRS), the present paper focusses on just the MODIS products owing to their widespread use in fire activity analyses in the US and elsewhere over the past decade (e.g. Urbanski *et al.* 2011; Hawbaker *et al.* 2013; Mouillot *et al.* 2014; Parks 2014; Schroeder *et al.* 2014).

Bias and uncertainty

The following sections are intended to provide an overview of the major reporting biases, inconsistencies, and sources of errors and uncertainty in the different types of US wildfire activity datasets (Fig. 1). It is beyond the scope of this paper to highlight and detail every issue that may concern a keen analyst. When topics have been covered in more depth elsewhere, readers will be referred to other published work for further information.
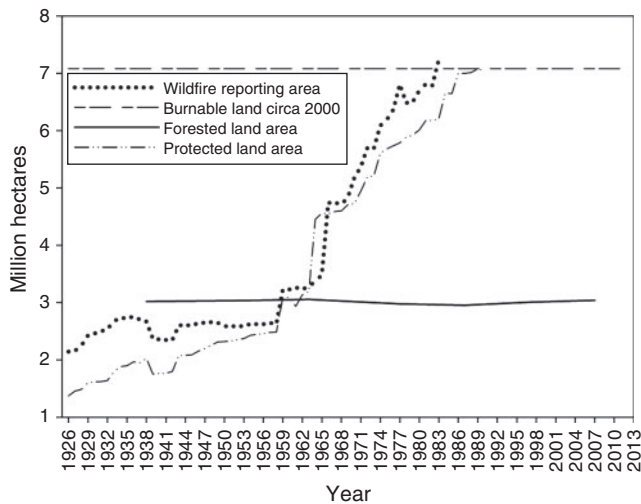
Archival summary reports

The USFS annual summary reports cover the longest time period, but are neither complete nor consistent over the ~80 years that they are available, because wildfire activity was tallied from an increasing land area over time. This changing baseline must be taken into account when analysing and interpreting the data. The reporting area more than triples in size from 1926 to 1983 (Fig. 2), which is the time period for which area-reporting figures are included in the USFS annual summaries. The reporting area rises as the land base qualifying for the federal fire protection program that required the reports grows from just the forested and other critical watersheds mandated by law in the earliest years to nearly the total burnable US land area by the early 1980s (Fig. 2). (The estimate of total burnable wildland in the US used here is based on data generated by the national LANDFIRE program using ~2000 satellite imagery (Reeves *et al.* 2006)). Even then, not all of the reporting area was formally protected from wildfire under the program requiring the statistics. Therefore, the USFS wildfire activity data are classified as from 'protected' and 'unprotected' areas, with unprotected areas having no organised fire protection despite their need for it as per the national directive at the time (explained further by Houghton *et al.* 2000). The total area of protected land estimated in the USFS reports does not reach total US burnable area until the last year it is included in the reports, which is 1990 (Fig. 2).

Because the USFS summary reports are broken down by US state, the reporting biases can be examined further at that resolution. Fig. 3 shows the percentage burnable area reporting, by contiguous US state, as an average within the three decades from 1950 to 1979, which correspond to a collective period of marked increase in land area accounted for in the USFS reports (see Fig. 2). By the 1970s, the reporting area begins to reach what could be considered an 'all-lands' coverage, but not in all states. Some states, like Alaska and Kansas, are completely unaccounted for in the USFS wildfire activity statistics before the 1940s and 1950s; others, like Texas and North Dakota, remain poorly accounted for well into the 1970s.

One reason for the large increase in land area accounted for in the USFS wildfire summary reports during the 1960s is the
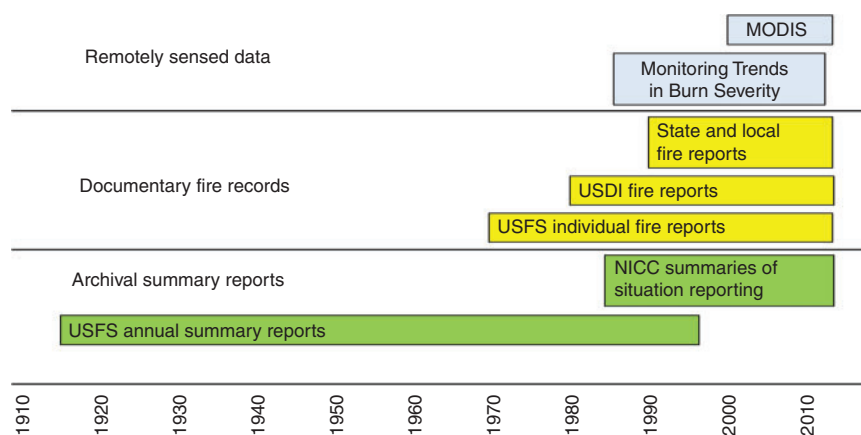

Fig. 1. Timeline showing the general temporal spans of the sources of US wildfire activity data discussed in this paper. MODIS, Moderate Resolution Imaging Spectroradiometer; USDI, US Department of Interior; USFS, US Forest Service; NICC, National Interagency Coordination Center.

Fig. 2. US wildfire reporting area (dotted line), 1926–83, and protected land area (dot-and-dash), 1926–90, from the US Forest Service (USFS) archival summary reports. The upper dashed reference line represents the ~700 million ha of US land, including Alaska, mapped as burnable from ~2000 satellite imagery by the national LANDFIRE program. The lower horizontal reference line represents ~300 million ha of US forested land estimated at various points in time during this period by the national Forest Inventory and Analysis program.

inclusion of an additional ~70 million ha of western US rangelands administered by the USDI Bureau of Land Management (BLM) (Fig. 4). This addition occurred after wildfires burned more than 100 000 hectares of BLM lands near Elko, Nevada, in 1964 and brought the BLM fully into the cooperative fire control program for which the USFS reports were generated (Pyne 1982). Estimates of wildfire activity levels on BLM lands in the contiguous US (CONUS) were first included in the 1966 USFS summary reports. Prior to 1966, the USFS statistics accounted for wildfire activity on only approximately half of the contemporary federal land base in the 11 westernmost states in the CONUS (Fig. 4). Littel *et al.* (2009) recognised the reporting bias and, as part of their analyses of fire–climate relationships in the western US, multiplied the USFS-reported WFAB estimates 'by the ratio of the total area protected in 2003 to the area protected in a given year'. Such adjustment assumes that wildfire activity levels are largely equivalent on a per-hectare basis between the reporting and non-reporting areas. However, the areas unaccounted for in the early USFS wildfire statistics (e.g. pre-1960s) for the US West should comprise a much greater proportion of non-forested land than the areas reporting, potentially invalidating any assumption of comparable fire-activity levels and possibly contributing to the underperformance of fire–climate models based on WFAB estimates
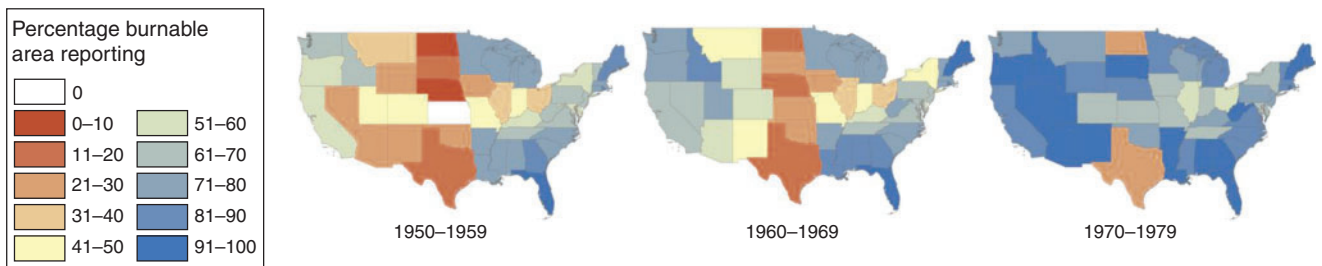
Fig. 3. Percentage of burnable land area accounted for in the US Forest Service (USFS) annual wildfire summary reports, as an average for each contiguous US state during the three decades from 1950 to 1979.
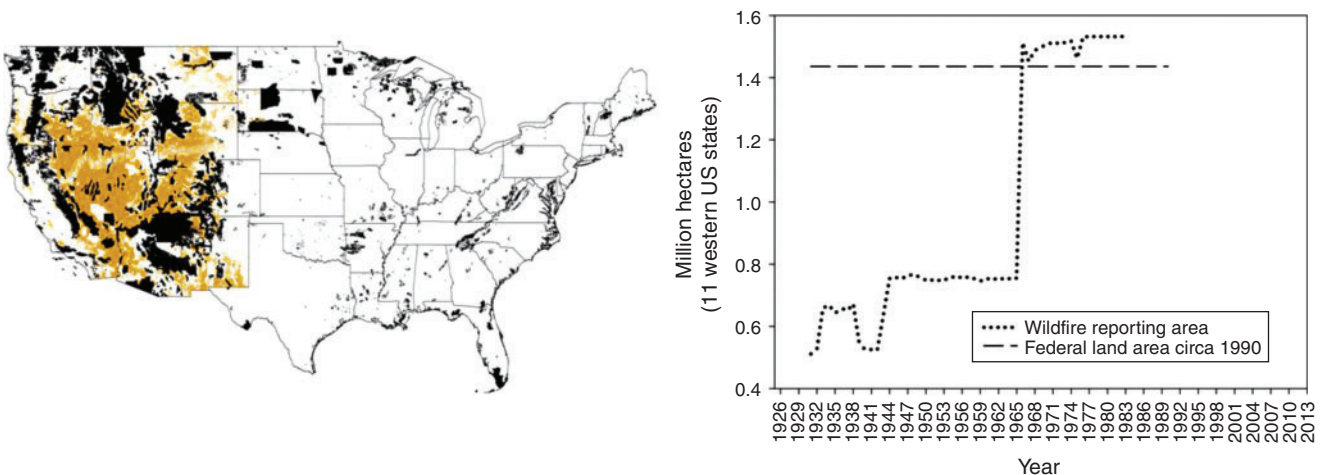
Fig. 4. Map of the contiguous US (CONUS) showing lands administered by the US Department of Interior Bureau of Land Management (USDI BLM) in orange and other federal lands in black. Wildfires on these BLM lands, which include ~70 million ha of largely non-forested area in the western US, are not accounted for in the US Forest Service (USFS) wildfire summary reports until 1966, after the BLM joined the cooperative fire control program for which the USFS reports were generated. The graph at right shows the spike in reporting area. The dashed reference line represents the ~140 million ha of total federal land in the 11 westernmost states of the CONUS as of the year 1990.

that included the early USFS figures (e.g. pre-1977) versus those that did not (Littel *et al.* 2009).

In addition to the area-reporting bias in the USFS summary reports, analysts should be aware of other inconsistency and uncertainty in the wildfire activity estimates that are included, especially for unprotected areas. Intentional ('controlled') burning was used extensively for vegetation management on non-federal lands, especially in the south-eastern US during the early 20th century. Although now used to a lesser extent (but on both federal and non-federal lands) in the US, intentional burning is not classified in the current reporting systems as 'wildfire' unless the controlled burn escapes and requires a suppression response. However, the early USFS wildfire activity summaries do include millions of hectares of intentional burning on 'unprotected' lands, which, until approximately the mid-20th century was viewed by the USFS as akin to wildfire, as something that should be prevented and ultimately eradicated (Pyne 1982). Controlled burning was accepted as a viable land-management practice over time and persists to this day (Melvin 2012); however, statistics regarding its use have not been included in summaries of 'wildfire' activity for several decades. Furthermore, the USFS includes with its unprotected area estimates the following caveat: 'since no field organisations are established in unprotected areas to report fires, the statistics on such unprotected areas are merely the best estimates by local agencies'. Although the degree of error in the reported figures cannot be known, gross errors apparently exist at least in some of the WFAB estimates from unprotected and protected areas. For example, the USFS annual summary reports for 1976–83 indicated that ~800 000 ha were burned by wildfires in Texas during each of 4 years within that 8-year period. Those are levels comparable with those witnessed in the state during what have been considered unprecedented fire seasons of 2006 and 2011 – seasons that prompted disaster declarations by the Federal Emergency Management Agency. There are multiple independent sources, including remotely sensed data, that corroborate the 2006 and 2011 estimates, but no records, including media reports, have been found to corroborate the similarly high USFS estimates from 1976 to 1983. Several authorities within the Texas state fire service have asserted that the USFS estimates must be in error, because the fire service has no recollection of large fires burning in 1976–83 on par with those in 2006 and 2011, and those recent 'outbreak' years were so significant that records of similar years, had they occurred just a few decades prior, would abound, especially in media archives (C. Stripling, T. Spencer, B. Smith, pers. comm.). If millions of hectares did not burn in Texas during each of several years in 1976–83, then any regional or national WFAB estimates based on the USFS summary reports are inflated as well. The total US WFAB reported by the USFS for each of the affected years is 1.6–2 million ha, with 35–50% based on the apparently inflated Texas unprotected-area estimates.

Potential errors in the wildfire activity statistics in the USFS annual summary reports are not limited to Texas, nor are they limited to figures from unprotected areas. Another potentially significant example is found in the USFS report for 1977. In that year, the USFS WFAB figure for protected lands in Alaska is 40 000 ha. However, WFAB estimates for that state in 1977 (~900 000 ha) published by Gabriel and Tande (1983) and

Hanson (1987) in conjunction with the 1977 Alaska fire season summary of McBride (1978) indicate that the USFS figure is an extreme underestimate.

Although the reporting biases and potential errors in the wildfire activity statistics in the USFS annual summary reports should be recognised and addressed in analyses for which they are used, a portion of the data have been made readily available online with no formal citation or any reference to their source. At the time of the present writing, the excerpt includes estimates of total wildfire numbers and area burned for the US from 1960 to 1982, and exists in tabular form on the website of the External Affairs section of the National Interagency Fire Center (NIFC), which is where the NICC is located. The NIFC webpage with figures for a total span of 1960–2014, http://www.nifc.gov/fireInfo/fireInfo_stats_totalFires.html (accessed 29 May 2015), is among the first returned in an internet search for 'wildfire statistics' or a similar term. Although the estimates are identified as 'total wildland fires and acres', at the time of the current writing, the figures represent wildfires only, and exclude intentional or controlled burning, which is a type of wildland fire by definition of the National Wildfire Coordinating Group (NWCG 2014). The figures for 1983–2014 are attributed to the Situation Reporting system that generates the estimates of wildfire numbers and area burned in the NICC annual reports described under 'Data sources' (see also Fig. 1). However, the same footnote explains that 'before 1983, sources of these figures are not known, or cannot be confirmed, and were not derived from the current situation reporting process. As a result the figures before 1983 shouldn't be compared with later data'. Yet, when the NIFC figures are plotted with the USFS annual summary reports (Fig. 5), whether as wildfire numbers or area burned, it is clear that NIFC External Affairs is using USFS estimates for 1960–82.
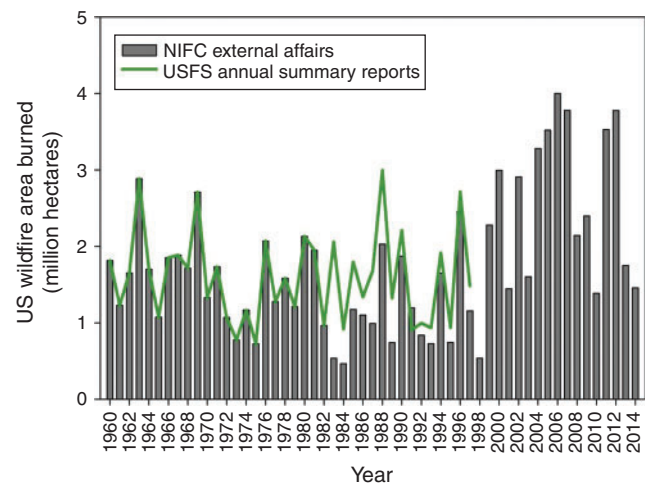
Fig. 5. Comparison of US wildfire area burned (WFAB) estimates distributed by National Interagency Fire Center (NIFC) External Affairs for 1960–2014 (bars) with estimates for 1960–97 included in the US Forest Service (USFS) archival summary reports (line). The figures, whether wildfire numbers or WFAB, are identical between the sources for 1960–82, indicating that the USFS summary reports are the source of the NIFC figures. The figures for 1983–97 differ between the sources, because, as NIFC explains, its estimates are from the National Interagency Coordination Center (NICC) summaries of the Situation Reporting data.

When the estimates from the USFS publications and the NIFC estimates sourced from the NICC Situation (SIT) Reporting summaries overlap in 1983–97, the USFS figures tend to meet or exceed the values in the NICC summaries. As noted above, the USFS estimate for US WFAB for 1983 may be inflated by as much as 800 000 ha. Although additional errors may exist in the USFS statistics, the NICC numbers may be biased downward during at least a portion of the period of overlap. Use of the SIT Reporting application, on which the NICC summary reports are based, is required only by agency and local dispatch offices associated with federal lands; it is effectively optional for non-federal units with wildfire protection responsibility only for non-federal lands. Voluntary SIT Reporting by non-federal units may result in considerable underestimation of elements like wildfire numbers and area burned in the NICC annual reports. The magnitude of the discrepancies between the USFS reports and the NICC numbers appears largest in the earliest years of SIT Reporting (~1983; Fig. 5), when the network of participating units was likely at its smallest.

Documentary fire records

In general, documentary fire records of the federal, state and local fire organisations are warehoused and available in electronic format only for a recent period of several decades (Fig. 1). The numerous US fire organisations have disparate wildfire reporting requirements, standards, protocols and archival systems, and compilation of data from multiple systems is difficult and time-consuming (Short 2014). It is therefore typical for wildfire activity analyses based on the documentary fire records to be limited to a single agency, like the USFS, or to federal lands only (e.g. Westerling *et al.* 2003, 2006).

Yet, even within a single fire organisation's system of record, the available electronic archive does not necessarily provide a complete or consistent accounting of actual wildfire activity. The USFS digital fire-reporting archive, for example, extends back to 1970. However, records from 1970 to 1992 were stored on computer tape before they were migrated, in 1992, to the current database archive (Bunton 2000). Bunton (2000) explains that, once the data from that early period were written to tape, no corrections could be made. Moreover, 'some fire reports were never added to the tape files, and . . . most of the missing fires were very large fires that were still active (not declared out) until well after the annual report for the year was completed' (Bunton 2000). Once they were migrated from tape to database, records from before 1986 remained uncorrectable, although Bunton (2000) indicates that some records of missing large fires were added after the current archival system was initiated in 1992.

Analysts commonly consider the USDI fire records complete only back to 1980, because the digital archive is erratically populated before that year. However, formal and comprehensive fire reporting was not consistently mandated by the USDI bureaus until the mid-1980s, and the USDI archive suffers from many of the data quality and completeness issues, as does that of the USFS before ~1992 (S. Larrabee, pers. comm.).

Each of the 50 US states experiences wildfires that occur outside of the protection responsibility areas of federal government, and each maintains records of those events in non-federal reporting systems. Although the period of record in the digital archives differs among the various non-federal fire organisations, the degree and quality of formal fire reporting has generally risen over time (Thomas and Butry 2012; Short 2014). For analyses of wildfire activity that include figures from states like Texas, for which wildfire protection and reporting are almost entirely the responsibility of the state or local service (see Artley 2009), it will be crucial to recognise and account for any non-federal reporting idiosyncrasies within the domain of interest; they can be substantial (Short 2014).

The content, quality and completeness of the available data can vary considerably among organisations and over time. For example, in the federal systems, the spatial resolution, or precision, of a fire's reported point of origin has generally improved over time from a legal description – township, range and section – to actual point coordinates in latitude and longitude. However, most local fire department records can be spatially resolved to county or zip code at best – often referencing the zip code of a responding fire station. Location information from the state reporting systems can span the whole spectrum from a latitude–longitude point of origin to a nominal state designation only. Fire cause and final fire size are other basic occurrence-related elements that tend to range in both specificity and accuracy, if reported at all (Prestemon *et al.* 2013). Estimates of fire size from locally (e.g. manually) mapped and satellite-derived fire perimeters (polygons) may differ from those included in the incident reports, and all may overestimate area burned by including unburned islands within the fire perimeter (Kolden and Weisberg 2007).

Remotely sensed data

The use of satellite-derived fire detection and mapping products in wildland fire science has been expanding in the US and across the globe in recent decades (Mouillot *et al.* 2014). These remotely sensed datasets have proved particularly valuable where ground data are lacking or otherwise incomplete (Mouillot *et al.* 2014). In the US, areas of limited ground-based documentary wildfire data may include lands for which fire protection is the responsibility of a local fire service that does not submit fire activity information to the (voluntary) national fire department reporting system. Millions of burnable hectares, including large fire-prone areas within several states in the south-central US (e.g. Kansas, Oklahoma and Texas) may fall into this category in any given year (Short 2014). However, the MODIS active fire and burned area products have a notably short history among the US wildfire activity datasets, dating back only to ~2000 (Fig. 1). To benefit from the enhanced fire detection capability of dual sensors (Hawbaker *et al.* 2008), analyses of MODIS products must be restricted to post 2003, which is the first full year that instruments were in use on both the Aqua and Terra satellites. By contrast, the current MTBS large-fire perimeter (and burn severity) dataset nominally spans a much lengthier period, 1984–2014.

At present, the 30-m resolution MTBS burn scar dataset affords a consistently generated ~30-year geospatial archive of large-fire occurrence across most of the US. Fires are mapped regardless of land ownership or agency affiliation, and therefore the MTBS dataset can be used for all-lands analyses of spatially

explicit large-fire activity in recent decades. Yet, although it is intended to represent all fires greater than 404 ha in the western US and those greater than 202 ha in the eastern US, the MTBS dataset does miss some fires (Howard *et al.* 2014). Landsat scene selection in the MTBS mapping process is directed by reports of large fires in a compilation of documentary fire records from federal and state wildland fire reporting systems (Eidenshink *et al.* 2007). Fires missing from the MTBS compilation of documentary records – and fires with records lacking adequate temporal or spatial information to point to a specific Landsat scene – are unlikely to be included in the MTBS perimeter dataset, unless the burn scars are discovered by chance in the process of mapping other fires. Fires may likewise be missed or incompletely mapped owing to poor scene quality (e.g. clouds, smoke, shadows), obstructive tree canopy over surface fire scars, patchy burning within pixels, mismatches in pre- and post-fire imagery related to sun angle and phenology, or rapid (<1 year) post-fire vegetation recovery (e.g. in some grassland and sprouting-shrub systems) (Kolden and Weisberg 2007; Kolden *et al.* 2012). Fires are also missing from the MTBS dataset owing to gaps in the Landsat archive (J. Lecker, pers. comm.). Gaps are particularly evident during the Landsat 'commercialisation' period, 1985–99, when the acquisition and archiving of images was largely driven by market demand (Tucker *et al.* 2004). Landsat data gaps are the reason, for example, that the MTBS dataset currently includes only one mapped fire (444 ha) in Alaska for 1996 and 1997 – years in which Alaska WFAB is estimated at 243 000 and 810 000 ha respectively by other sources (e.g. Short 2014).

The MTBS burned-area boundaries are visually interpreted and manually delineated from Landsat imagery and derivative change indices (Eidenshink *et al.* 2007). Although the MTBS data suite includes pixel-level severity estimates for mapped fires, the identified unburned areas are not excised by rote from fire perimeters included in the polygon-based MTBS product. The unburned portions within the MTBS (and other) mapped perimeters can be substantial. As of this writing, 18.5% of the US wildfire area mapped by MTBS, 1984–2014, falls within the 'unburned to low' severity category (B. Quayle, pers. comm.). That figure is consistent with estimates of unburned areas within fire perimeters from local studies. Kolden *et al.* (2012) reported a total of 14–35% unburned area within fire perimeters from three national parks in the western US, 1984–2009, mapped using the MTBS protocol. Kasischke and Hoy (2012) found a similar average of 20% unburned area within MTBS fire perimeters from Alaskan forest fires in 2004 and 2006–08. However, the unburned fraction can vary greatly among fires, and has been found to be inversely related to factors like fire size and severity (Kolden *et al.* 2012). If the size and severity of wildfires, for example, increase over time, the unburned portions within fire polygons may decrease. Kolden and Weisberg (2007) describe in further detail the challenges and associated sources of error in the mapping and measuring of WFAB, both locally and remotely via the MTBS protocol, and readers are referred to that publication for more on these topics.

The MTBS dataset identifies prescribed (or controlled) burns separately from wildfires. This 'fire type' attribute is useful for screening out the intentional, controlled burns for an analysis of wildfire activity. However, before the April 2014 product release, fires mapped by MTBS defaulted to a fire-type label of 'wildfire', and that default label was updated to another type (e.g. prescribed fire) only when the associated fire report specified it accordingly. For this reason, all fires mapped without the aid of documentary records (e.g. fires discovered in the process of mapping others) were identified as wildfires. In the April 2014 release, this issue was remedied to some degree by relabelling any fires of uncertain fire type to 'unknown'. Of the ~24 000 fire polygons now in the April 2015 MTBS product release (the most recent as of the present writing), 4546 (19%) are labelled as unknown fire type. Analysts must determine the appropriate use of these uncertain types. Many of the fires of unknown type map within the Flint Hills region of Kansas and northern Oklahoma, where ~700 000 ha are intentionally burned each spring for rangeland improvement (KDHE 2010; Melvin 2012). Fires of unknown type in that region are prime candidates for exclusion from MTBS-based analyses of wildfire activity.

Improvements in spaceborne sensors over the past couple of decades have expanded the available set of data products characterising fire activity and associated burned area (Mouillot *et al.* 2014). Products derived from MODIS sensors are increasingly used for analyses of wildland fire activity, as the instruments employ what are considered to be the most useful channels (spectral bands) for fire monitoring and mapping – at spatial resolutions of 500 m (burned area) to 1 km (active fire). The MODIS active fire product (Giglio *et al.* 2003) is based on detection of thermal anomalies, and although strictly a fire detection product, it has been used to estimate burned area in the US (e.g. Wiedinmyer *et al.* 2011; Hawbaker *et al.* 2013). Temporal sampling by the MODIS sensors (on two Earth Observation Satellites) is limited: for most of CONUS, a total of four satellite overpasses comprise a midday pair and a nighttime pair. Adjacent MODIS overpasses by an individual sensor increasingly overlap at higher latitudes, and locations in boreal and Arctic regions of Alaska, for example, may be observed up to four times per sensor per 24-h period (Ambrosia *et al.* 2014). Nevertheless, the sensors may not pass over as burning occurs, and clouds (or smoke) may preclude active fire detection during the limited overpasses (Giglio *et al.* 2006; Hawbaker *et al.* 2008; Urbanski *et al.* 2009). In general, the MODIS active fire product is subject to high omission errors with regard to fires of small size, short duration or low intensity and may underestimate area burned by large, rapidly moving fires that travel across multiple MODIS pixels between overpasses (Hawbaker *et al.* 2008). The MODIS sensors are generally considered to have a lower fire-size detection threshold of $\sim 100\,\mathrm{m}^2$ under the most favourable conditions (Giglio *et al.* 2003). But above that size threshold, omission errors can be quite high and tend to be much greater in grasslands and open shrublands than in forested areas (Hawbaker *et al.* 2008; Urbanski *et al.* 2009; Mouillot *et al.* 2014). According to Urbanski *et al.* (2009), 'even large rangeland fires (>2000 ha) may completely evade MODIS detection'. Within a given vegetation type, however, active-fire omission errors tend to decrease with increasing fire size (Hawbaker *et al.* 2008; Urbanski *et al.* 2009).

Errors of commission (i.e. false positives) in the MODIS active fire product may result from confusion of highly reflective non-fire surfaces with true fires, or, owing to its contextual

algorithm, from sharp radiometric contrasts between adjacent non-fire elements, like certain vegetation and soils (Hawbaker *et al.* 2008). More importantly and despite omission errors, burned areas derived from MODIS active detection products tend to overestimate the true burned area because the sensor's detection threshold of $\sim$100 m^2 is much lower than the product's 1-km pixel resolution (Giglio *et al.* 2006). Thus, although the MODIS active fire product may capture the majority of certain types of large fire events (e.g. forest fires), the corresponding burned area estimates are highly error-prone. Adjustments to initial burned-area estimates, including scaling by fraction of vegetation cover or a combination of vegetation cover and fire-pixel clustering, may improve burned-area estimates derived from the active-fire product (Giglio *et al.* 2006; Wiedinmyer *et al.* 2011).

Aside from the thermal-based active-fire product, several 500-m-resolution MODIS burned-area products based on surface reflectance (burn scars) have been generated (Mouillot *et al.* 2014). Roy *et al.* (2005) and Giglio *et al.* (2009) identify burn scars based on changes in surface reflectance, whereas Urbanski *et al.* (2009) developed an algorithm to produce near-real-time burn scar data using a single observation of surface reflectance. The burn scar algorithm of Roy *et al.* (2005) is employed in the standard, monthly MODIS burned area product distributed by the US NASA. Both Giglio *et al.* (2009) and Urbanski *et al.* (2009) leverage the active fire detections in their algorithms to confirm burn scars and reduce uncertainty in burned area estimates; using a version of this method (and MTBS burned pixels rather than NICC burned area estimates as 'ground-truth' data), Urbanski *et al.* (2011) were able to reduce the apparent upward bias in MODIS burned-area estimates for the western US, 2003–08, to 7% and reported uncertainty in their annual estimates of area burned by wildland fire to be $\leq$5%.

From a *wildfire* activity standpoint, another important source of uncertainty in the MODIS-based products is the conflation of wildfire and intentional, controlled (or prescribed) burning in both wildland and agricultural areas. All of the controlled burning must be screened from remote-sensing fire products for any analyses of wildfire activity, per se. According to estimates from the NICC, an average 800 000 ha of wildlands have been intentionally burned (with prescribed fire) in the US annually since the NICC began keeping records in 1998. However, that figure is a conservative estimate, because much controlled burning on private lands, including commercial timberlands, is not reported to the NICC. An independent survey of controlled burning levels for the year 2011 by the National Association of State Foresters and the Coalition of Prescribed Fire Councils, suggests that the total area deliberately burned in the US during that year was 8 million ha, or *10 times* the NICC estimate (Melvin 2012). Of that total, 39% (3.1 million ha) was reported as controlled burning on forested lands, predominantly in the south-eastern US, whereas the remainder was considered agricultural burning, largely in western and central plains states (Melvin 2012). Assuming the 2011 controlled burning figures are typical, which Melvin (2012) suggests, and given that the average annual area burned by wildfire over the past 20 years is estimated (from several sources) as $\sim$2.3 million ha (Short 2014), then most of area burned annually in the US is not from

wildfire, but from intentional burning for agricultural or other land-management purposes. Moreover, the contemporary annual area burned by wildfire in the US appears likely to be less than that burned deliberately for forestry objectives. Therefore, even if MODIS detection rates from agricultural areas are expected to be fairly low (Hawbaker *et al.* 2008) and even if remotely sensed landcover maps can be used to exclude detections in pixels mapped as developed or agriculture or cropland (McCarty *et al.* 2007; Urbanski *et al.* 2011; Hawbaker *et al.* 2013), additional efforts must be made to exclude controlled burning in forested areas. Otherwise, 'wildfire' activity levels would likely be greatly overestimated from MODIS products, especially if the south-eastern US is included in the analysis.

Mouillot *et al.* (2014) expand on many of the points made here with regard to potential limitations of the MODIS (and other satellite-based) products and discuss them specifically in light of potential uses of the data, including for hazard assessment, mobilisation of firefighting resources and emissions modelling; readers seeking more information on these topics are encouraged to consult that publication. With growing interest in MODIS-based active-fire count and burned-area datasets, there is increasing pressure to improve the accuracy, resolution and temporal specificity of the products (Mouillot *et al.* 2014; Schroeder *et al.* 2014). Any changes in detection capabilities or processing algorithms that result in such improvements must be taken into account in any multiannual analysis to potentially mitigate influences of temporal instability in the datasets.

Summary and conclusions

There is a wealth of US wildfire activity data available for analyses today, but analysts must be aware of inherent reporting biases, inconsistencies and uncertainty in the data in order to maximise the integrity and utility of their work. Users of data from archival summary reports should recognise that the estimates come from an increasing land area over time, and even the most recent annual wildfire summary reports from the NICC do not account for all wildfire activity (especially fire numbers) in the US, owing in part to limited wildfire reporting by local fire departments (Thomas and Butry 2012). Moreover, all estimates are just that, and area-burned estimates are inherently uncertain. Burned-area estimates based on fire perimeters, which often include significant unburned portions, are typically biased high ($\sim$15–30%). Analysts attempting to assess MODIS burned-area products have, for example, adjusted the NICC reference figures of area burned by a factor of 0.72 to account for a perceived overestimation of 28% (Urbanski *et al.* 2009). Unadjusted, area-burned figures based solely on the MODIS active-fire product tend to be overestimates, due to the ability of the sensors to detect fires smaller than the products' pixel sizes. Algorithms that use the active fire detections to confirm burn scars can generate MODIS burned area products that align well with those estimated from MTBS burn scars (with unburned islands excluded).

All remotely sensed datasets intended for analyses of wildfire activity, per se, must be carefully screened to remove areas burned intentionally in both agricultural areas and wildlands, which may account for 80% of the total area burned in the US in any given year. Millions of hectares of intentional burning in the

south-eastern US are included as part of the annual estimates of wildfire activity on unprotected lands for several early decades of the USFS annual summary reports. It is inappropriate to compare, for example, the early total area burned estimates, which factor in 'prescribed fire', with the statistics in later USFS reports, which omit it, or to compare them with figures post 1998 in the NICC reports, which greatly underestimate levels of intentional burning in the US.

Documentary fire records from the digital archives of federal wildfire reporting are unlikely to represent the total wildfire activity levels on federal lands in years before 1992, but it is unclear how much data is missing. Use of the national fire reporting systems of state and local fire organisations has been rising in recent decades (Thomas and Butry 2012; Short 2014), providing an improved set of incident-level data for all-lands analyses of wildfire activity. Short (2014) expanded on earlier efforts of Schmidt *et al.* (2002) to compile federal and non-federal wildfire data from the various reporting systems for the national Fire Program Analysis (FPA) system. The resulting FPA Fire-Occurrence Database (FPA FOD) currently includes 1.73 million records from the 22-year period 1992–2013, with values for at least the following core data elements: location (fire origin) at least as precise as a Public Land Survey System section ($2.6\text{-}km^2$ grid), discovery date and final fire size. The FPA FOD is publicly available from the USFS Research Data Archive (http://dx.doi.org/10.2737/RDS-2013-0009.3, accessed 29 May 2015). The FPA FOD is based on geospatial documentary fire records, but has been evaluated for completeness using archival summary reports, including those published by the NICC (Short 2014). Although unavoidably incomplete in some aspects (Short 2014), the FPA FOD should facilitate fairly high-resolution geospatial analysis of US wildfire activity over recent decades, using data from multiple systems of record. Additional elements added to the original fire records, including the unique MTBS perimeter identifiers, effectively provide a 'bridge' between the ground- and satellite-based datasets, and the MTBS perimeters can be used in conjunction with the FPA FOD to expand on the geospatial analyses possible with point-of-origin information in the original fire reports (Short 2014). Analysts are encouraged to consider the FPA FOD as a source of integrated US wildfire activity data from recent decades. Short (2014) provides further guidance in this regard.

Acknowledgements

References

Ambrosia VG, Giglio L, Schroeder W, Csiszar I, Quayle B (2014) Satellite and airborne fire sensor systems for Arctic wildfire observations. A report for the Interagency Arctic Research and Policy Committee: Wildfire Implementation Team, V5 (11–7-14). Available at http://www.iarpccollaborations.org/uploads/cms/documents/wildfire-sensor-systems_v5.pdf [Verified 29 May 2015]

Artley DK (2009) Wildland fire protection and response in the United States: the responsibilities, authorities, and roles of federal, state, local, and tribal government. Report for the International Association of Fire Chiefs. Available at https://www.iafc.org/files/wild_MissionsProject.pdf [Verified 29 May 2015]

Bunton DR (2000) Wildland fire and weather information data warehouse. In 'Proceedings of the seventh symposium on systems analysis on forest resources', 28–31 May 1997, Traverse City, MI. USDA Forest Service, North Central Forest Experiment Station, General Technical Report GTR-NC-205, pp. 297–302. (Saint Paul, MN)

Collins BM, Omi PN, Chapman PL (2006) Regional relationships between climate and wildfire-burned area in the Interior West, USA. *Canadian Journal of Forest Research* **36**, 699–709. doi:10.1139/X05-264

Dennison PE, Brewer SC, Arnold JD, Mortiz MA (2014) Large wildfire trends in the western United States, 1984–2011. *Geophysical Research Letters* **41**, 2928–2933. doi:10.1002/2014GL059576

Eidenshink J, Schwind B, Brewer K, Zhu Z, Quayle B, Howard S (2007) A project for monitoring trends in burn severity. *Fire Ecology* **3**, 3–21. doi:10.4996/FIREECOLOGY.0301003

Gabriel HW, Tande GF (1983) A regional approach to fire history in Alaska. USDI Bureau of Land Management, BLM-Alaska Technical Report No. 9. (Anchorage, AK)

Giglio L, Descloitres J, Justice CO, Kaufman YJ (2003) An enhanced contextual fire detection algorithm for MODIS. *Remote Sensing of Environment* **87**, 273–282. doi:10.1016/S0034-4257(03)00184-6

Giglio L, van der Werf GR, Randerson JT, Collatz GJ, Kasibhatla P (2006) Global estimation of burned area using MODIS active fire observations. *Atmospheric Chemistry and Physics* **6**, 957–974. doi:10.5194/ACP-6-957-2006

Giglio L, Loboda T, Roy DP, Quayle B, Justice CO (2009) An active-fire based burned area mapping algorithm for the MODIS sensor. *Remote Sensing of Environment* **113**, 408–420. doi:10.1016/J.RSE.2008.10.006

Hanson, RE (1987) Historical fire data – BLM Alaska: 1959–1985. Bureau of Land Management Alaska Open File Report 22. (Anchorage, AK)

Hawbaker TJ, Radeloff VC, Syphard AD, Zhu Z, Stewart SI (2008) Detection rates of the MODIS active fire product in the United States. *Remote Sensing of Environment* **112**, 2656–2664. doi:10.1016/J.RSE.2007.12.008

Hawbaker TJ, Radeloff VC, Stewart SI, Hammer RB, Keuler NS, Clayton MK (2013) Human and biophysical influences on fire occurrence in the United States. *Ecological Applications* **23**, 565–582. doi:10.1890/12-1816.1

Houghton RA, Hackler JL, Lawrence KT (2000) Changes in terrestrial carbon storage in the United States. 2: The role of fire and fire management. *Global Ecology and Biogeography* **9**, 145–170. doi:10.1046/J.1365-2699.2000.00164.X

Howard SM, Picotte JJ, Coan MJ (2014) Utilizing multi-sensor fire detections to map fires in the United States. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **XL-1**, 161–166. doi:10.5194/ISPRSARCHIVES-XL-1-161-2014

KDHE Kansas Department of Health and Environment (2010) Flint Hills smoke management plan. A report of the State of Kansas. Available at http://www.ksfire.org/docs/about/Flint_Hills_SMP_v10FINAL.pdf [Verified 29 May 2015]

Kasischke ES, Hoy EE (2012) Controls on carbon consumption during Alaskan wildfires. *Global Change Biology* **18**, 685–699. doi:10.1111/J.1365-2486.2011.02573.X

Kolden CA, Weisberg PJ (2007) Assessing accuracy of manually mapped wildfire perimeters in topographically dissected areas. *Fire Ecology* **3**(Special issue), 22–31. doi:10.4996/FIREECOLOGY.0301022

Kolden CA, Lutz JA, Key CH, Kane JT, van Wagtendonk JW (2012) Mapped versus actual burned area within wildfire perimeters: characterizing the unburned. *Forest Ecology and Management* **286**, 38–47. doi:10.1016/J.FORECO.2012.08.020

Lannom KO, Tinkham WT, Smith AMS, Abatzoglu J, Newingham BA, Hall TE, Morgan P, Strand EK, Paveglio TB, Anderson JW, Sparks

AM (2014) Defining extreme wildland fires using geospatial and ancillary metrics. *International Journal of Wildland Fire* **23**, 322–337. doi:10.1071/WF13065

Littel JS, McKenzie D, Peterson DL, Westerling AL (2009) Climate and wildfire area burned in western US ecoprovinces, 1916–2003. *Ecological Applications* **19**, 1003–1021. doi:10.1890/07-1183.1

McBride FE (1978) Alaska fire season – 1977. *Fire Management Notes* **39**, 3–7.

McCarty JL, Justice CO, Korontzi S (2007) Agricultural burning in the south-eastern United States detected by MODIS. *Remote Sensing of Environment* **108**, 151–162. doi:10.1016/J.RSE.2006.03.020

Melvin M (2012) National prescribed fire use survey report. Technical Report 01–12, Coalition of Prescribed Fire Councils, Inc. Available at http://www.stateforesters.org/sites/default/files/publication-documents/2012_National_Prescribed_Fire_Survey.pdf [Verified 29 May 2015]

Mouillot F, Field CB (2005) Fire history and the global carbon budget: a 1° × 1° fire history reconstruction for the 20th century. *Global Change Biology* **11**, 398–420. doi:10.1111/J.1365-2486.2005.00920.X

Mouillot F, Schultz MG, Yue C, Cadule P, Tansey K, Ciais P, Chuvieco E (2014) Ten years of global burned area products from spaceborne sensing – a review: analysis of user needs and recommendations for future developments. *International Journal of Applied Earth Observation and Geoinformation* **26**, 64–79. doi:10.1016/J.JAG.2013.05.014

NWCG National Wildfire Coordinating Group (2014) Glossary of wildland fire terminology. NWCG Publication PMS-205. (Boise, ID) Available at http://www.nwcg.gov/pms/pubs/glossary/pms205.pdf [Verified 29 May 2015]

Parks SA (2014) Mapping day-of-burning with coarse resolution satellite fire-detection data. *International Journal of Wildland Fire* **23**, 215–223. doi:10.1071/WF13138

Prestemon JP, Hawbaker TJ, Bowden M, Carpenter J, Brooks MT, Abt KL, Sutphen R, Scranton S (2013) Wildfire ignitions: a review of the science and recommendations for empirical modeling. USDA Forest Service, Southern Research Station, General Technical Report SRS-171. (Asheville, NC)

Pyne SJ (1982) 'Fire in America: a cultural history of wildland and rural fire.' (University of Washington Press: Seattle)

Reeves MC, Kost JR, Ryan KC (2006) Fuels products of the LANDFIRE project. In 'Fuels management – how to measure success: conference proceedings', 28–30 March 2006, Portland, OR. (Eds PL Andrews, BW Butler) pp. 239–252. USDA Forest Service, Rocky Mountain Research Station, Proceedings RMRS-P-41. (Fort Collins, CO)

Riley KL, Abatzoglou JT, Grenfell IC, Klene AE, Heinsch FA (2013) The relationship of large-fire occurrence with drought and fire danger indices in the western USA, 1984–2008: the role of temporal scale. *International Journal of Wildland Fire* **22**, 894–909. doi:10.1071/WF12149

Roy DP, Jin Y, Lewis PE, Justice CO (2005) Prototyping a global algorithm for systematic fire-affected area mapping using MODIS time series data *Remote Sensing of Environment* **97**, 137–162. doi:10.1016/J.RSE.2005.04.007

Schmidt KM, Menakis JP, Hardy CC, Hann WJ, Bunnel DL (2002) Development of coarse-scale spatial data for wildland fire and fuel management. USDA Forest Service, Rocky Mountain Research Station, General Technical Report RMRS-87. (Fort Collins, CO)

Schroeder W, Oliva P, Giglio L, Csiszar IA (2014) The new VIIRS 375-m active fire detection product: algorithm description and initial assessment. *Remote Sensing of Environment* **143**, 85–96. doi:10.1016/J.RSE.2013.12.008

Short KC (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

Thomas DS, Butry DT (2012) Wildland fires within municipal jurisdictions. *Journal of Forestry* **110**, 34–41. doi:10.5849/JOF.10-024

Tucker CJ, Grant DM, Dykstra JD (2004) NASA's global orthorectified Landsat data set. *Photogrammetric Engineering and Remote Sensing* **70**, 313–322. doi:10.14358/PERS.70.3.313

Urbanski SP, Salmon JM, Nordgren BL, Hao WM (2009) A MODIS direct broadcast algorithm for mapping wildfire burned area in the western United States. *Remote Sensing of Environment* **113**, 2511–2526. doi:10.1016/J.RSE.2009.07.007

Urbanski SP, Hao WM, Nordgren B (2011) The wildland fire emission inventory: western United States emission estimates and an evaluation of uncertainty. *Atmospheric Chemistry and Physics* **11**, 12973–13000. doi:10.5194/ACP-11-12973-2011

Westerling AL, Gershunov A, Brown TJ, Cayan DR, Dettinger MD (2003) Climate and wildfire in the western United States. *Bulletin of the American Meteorological Society* **84**, 595–604. doi:10.1175/BAMS-84-5-595

Westerling AL, Hidalgo HG, Cayan DR, Swetnam TW (2006) Warming and earlier spring increase western US forest wildfire activity. *Science* **313**, 940–943. doi:10.1126/SCIENCE.1128834

Wiedinmyer C, Akagi SK, Yokelson RJ, Emmons LK, Al-Saadi JA, Orlando JJ, Soja AJ (2011) The fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning. *Geoscientific Model Development* **4**, 625–641. doi:10.5194/GMD-4-625-2011