Heliyon 6 (2020) e04159

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Spatial heterogeneity of winds during Santa Ana and non-Santa Ana wildfires in Southern California with implications for fire risk modeling

Alex W. Dye [a,*], John B. Kim [b], Karin L. Riley [c]

[a] Corvallis Forestry Sciences Laboratory, USDA Forest Service Pacific Northwest Research Station, ORISE Fellow, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
[b] Corvallis Forestry Sciences Laboratory, USDA Forest Service Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
[c] Fire Sciences Laboratory, USDA Forest Service Rocky Mountain Research Station, 5775 US Highway 10 W, Missoula, MT 59808, USA

ARTICLE INFO

Keywords:
Wildfire
Fire modeling
Southern California
Santa Ana winds
Remote access weather stations
Meteorology
Environmental analysis
Environmental hazard
Physical geography
Natural hazard
Environmental science
Geography

ABSTRACT

In Southern California, the Santa Ana winds are famous for their role in spreading large wildfires during the fall/winter season. Combined with Southern California's complex topography, Santa Anas create challenges for modeling wind-fire relationships in this region. Here, we assess heterogeneity of winds during Santa Ana and non-Santa Ana days, on days with and without large-fire ignitions, across a modern high-density observational network of 30 meteorological stations. Wind speeds on Santa Ana days with a large fire ignition (mean windspeed = 5.19 m/s) are significantly higher than on Santa Ana days without large fire ignitions (3.96 m/s), while on non-Santa Ana days winds are generally weaker, during both fire (2.30 m/s) and non-fire (2.38 m/s) days. Hierarchical clustering of meteorological stations during both Santa Ana and non-Santa Ana days reveals groups of stations with consistently similar wind speed and directions. All stations clearly exhibit high wind speeds on Santa Ana days, and most record contrasting wind characteristics during Santa Ana versus non-Santa Ana ignitions. Additionally, our analysis revealed that key geographic siting traits are not represented in the network, including few stations with northwest aspect and upper slope in the southern mountains.

1. Introduction

Santa Ana winds are an integral component of the Southern California climate and fire regime because of their ability to spread large wildfires [1]. Despite the natural adaptation of Southern California's vegetation to fire [2], Santa Ana fires cause billions of dollars in property damage [3], respiratory injury from smoke inhalation [4, 5], and enormous fire management expenditures [6, 7] due to proximity of population centers to burnable wildland areas. Future impacts of Santa Ana wildfires remain uncertain, but climate change and population growth will undoubtedly continue to increase socio-ecological stress during these events [7, 8, 9, 10]. Because Santa Ana winds are a critical component of the natural Southern California fire regime, understanding the spatial heterogeneity of winds on days when fires ignite, both during Santa Ana and non-Santa Ana conditions, is critical. The focus of our paper is to investigate this heterogeneity using a modern high-density network of observational meteorological data for the purpose of improving current and future representation of winds in wildfire simulation models.

Regional characteristics of Santa Anas have been studied using both synoptic climatological approaches [11, 12] and dynamically downscaled datasets [13, 14, 15]. Typically, Santa Ana winds occur between September and April, with higher likelihood from October to December [12, 15, 16]. High wind speeds, northeasterly wind directions, and low relative humidity coupled with a strong pressure gradient between the Great Basin and coastal California are meteorological indicators of Santa Ana winds [10, 11, 12, 15, 17, 18, 19, 20]. Although ignitions not associated with Santa Ana winds (usually in the summer) are more frequent, Santa Ana fires burn significantly more area [21]. Between 2001-2007, Santa Ana fires burned 3.5–4.5 times more area than non-Santa Ana fires [19]. Similarly, while just 24% of fires are related to Santa Ana winds, they can comprise up to 50% of total annual burned area [20]. Geographically, fires driven by Santa Ana winds are spatially distinct from those not driven by Santa Ana winds, igniting in less populated, high elevation areas, downwind from mountain slopes and decreasing in prevalence in desert basins and towards the ocean [7, 13, 22, 23].

The complex topography of Southern California creates scenarios where winds can enter eddies or funnel into mountain passes and canyons, forcing local variations in Santa Ana wind flow [24, 25, 26]. However, landscape variation of observed meteorological records during

* Corresponding author.
E-mail address: alex.dye@usda.gov (A.W. Dye).

https://doi.org/10.1016/j.heliyon.2020.e04159
Received 10 September 2019; Received in revised form 13 December 2019; Accepted 5 June 2020

Santa Ana wind days remains sparsely examined and is not always considered fully when modeling the effects of wind on fire spread, despite the vast quantities of resources devoted to management of fires driven by Santa Ana winds. Local wind behavior has obvious importance for modeling fire spread at specific locations on a landscape, particularly when observed winds are used as key inputs to fire spread models that actively inform fire mitigation spending and risk assessment [27, 28, 29, 30].

Many if not most fire simulation programs utilize weather drawn from a single weather station or scenario out of computational necessity (e.g. FlamMap [31], FARSITE [32], FSPro [33], FSim [34], and HFire [27]. However, some contain features that use station observations to create a wind field reflecting variability of wind speed and direction as it is forced over complex terrain (e.g. the WindNinja program that can be activated in FARSITE [32, 35]. In either case, careful station selection is necessary to produce valid outputs for fire prediction. Choosing a station that is affected by topography rather than synoptic conditions is unlikely to capture the wind regime of the area under study. Even nearby stations in different topographic settings have been shown to record different wind characteristics [36]. Thus, understanding the spatial variability in winds and how this might or might not be reflected by individual stations is necessary to producing outputs reflective of the area's fire regime.

The most detailed analysis of spatial variability of observed surface winds during Santa Ana days was conducted over 50 years ago [24]. Using observational hourly wind records from meteorological stations in the Los Angeles basin on 149 Santa Ana wind days, the authors found that winds generally fell into one of three categories: strong winds from the northeast, strong winds from the north, and weaker, generally northern, winds. This work has since been extended with more recent spatial and temporal modeling of Santa Ana wind variability [13, 15, 26, 37]. The next steps forward beyond these studies require examination of fine-scale wind heterogeneity using the modern high-density observational network that now exists across Southern California.

In our paper, we use this modern observational network to revisit ideas presented by [24] for the two-fold purpose of a) better grasping wind heterogeneity during fires linked to both Santa Ana and non-Santa Ana winds; and b) providing a reference for fire modelers to inform the implications of wind data selection in Southern California. Specifically, we ask the following question: how do observed hourly wind directions and speeds recorded at Southern California meteorological stations vary across the four categories of days, and what implications does station heterogeneity have for fire spread simulation modeling? This research is intended to aid ongoing efforts to better understand and model fire risk in complicated environmental systems like Southern California.

2. Materials and methods

2.1. Study area

We used Pyrome 34 (Figure 1) from a dataset crafted for national-level fire risk assessments [38]. Pyromes are regions characterized by a particular fire regime, as defined by fire size, frequency, intensity, and season [39, 40]. Analyses at the spatial scale of pyrome boundaries are commonly used in national fire modeling applications [38].

Pyrome 34 encompasses most of southwestern California, including the metropolitan areas of Los Angeles and San Diego and the surrounding mountain ranges and national forest lands (Figure 1). This region has a wet-winter/dry-summer Mediterranean climate, and vegetation consists

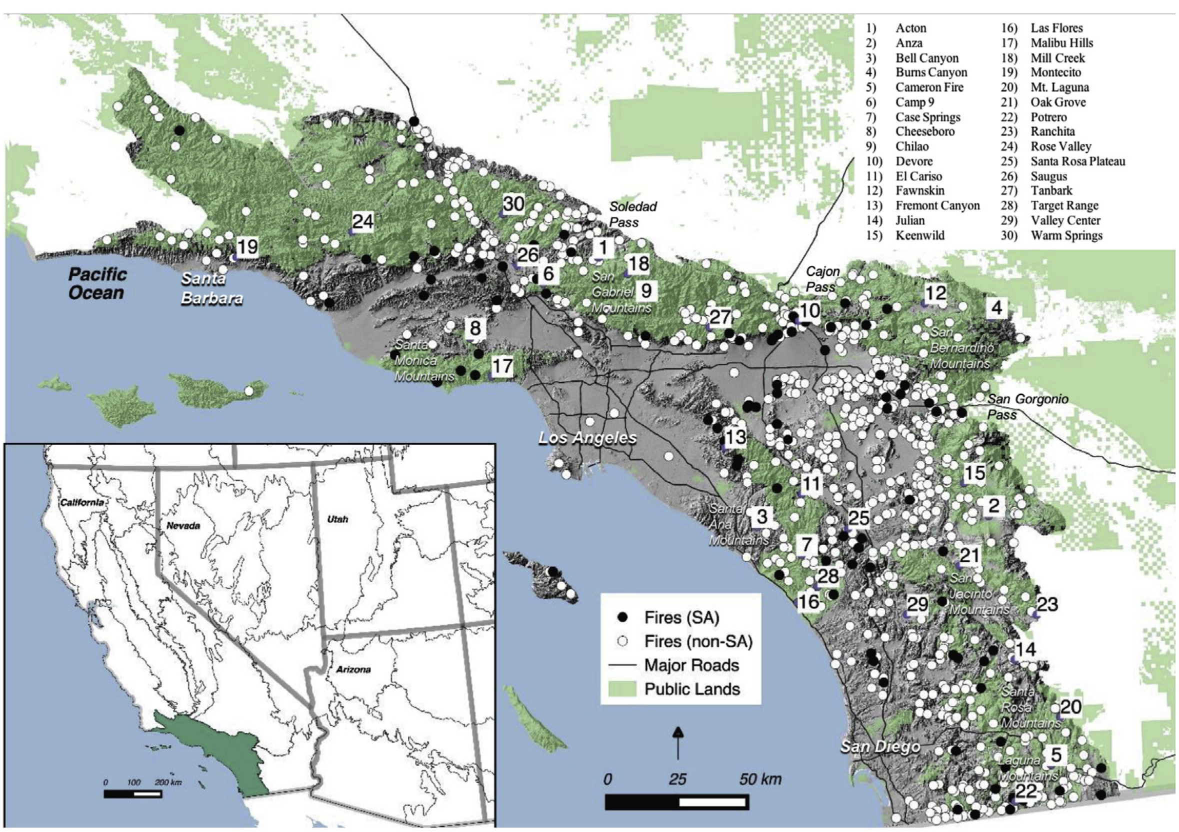

Figure 1. Topographic relief map of the Pyrome 34 study area in southwestern California. Locations of the 30 Remote Access Weather Stations (RAWS) are identified on the legend key. Dots indicate ignition location of large fires ignited during the period 1996–2010 on Santa Ana days (black) and non-Santa Ana days (white). Public lands are shaded in green; major roads are indicated by black lines. In the inset map, national USFS pyrome boundaries are outlined, with Pyrome 34 shaded green. Geographic features mentioned in the text are included for reference.

primarily of schlerophyllic woody evergreen shrub cover and oak woodland savannah, with pine, fir, and pinyon-juniper in the mountains. Topography is varied, ranging from sea level on the Pacific coast to mountain peaks exceeding 3,000 m. Southern California is home to three of the United States' most populous metropolitan areas (Los Angeles-Long Beach-Anaheim, Riverside-San Bernardino, and San Diego), with a combined population exceeding 22 million people. Fires are the primary ecological disturbance in the chaparral vegetation of California's Mediterranean climate [2, 9] and are typically driven by drought during the summer, and by high-speed Santa Ana winds in fall and winter.

2.2. Hourly weather station data

We obtained hourly 10-minute average observational wind data recorded at Remote Access Weather Stations (RAWS) [41]. RAWS are sited to record fire weather and have been extensively used in fire modeling applications throughout the western United States [27, 34, 42, 43]. For this reason, we only analyze the RAWS station network, and do not include other networks or models. Wind observations at all RAWS are recorded at a height of 6 m [41]. At this height, RAWS are able to measure the synoptic wind, unaffected by drag from the ground; fire models then calculate the estimated wind that would be on the fire based on vegetation and topographic characteristics.

We selected stations that met the following two criteria: 1) located within the Pyrome 34 boundary; and 2) consistent data coverage with <10% missing data extending back to 1996. This resulted in 30 stations from which we extracted hourly surface wind direction and wind speed (Figure 1). We chose 1996 as a starting point because RAWS data availability swiftly declines prior to 1996 and this allowed us to balance both temporal and spatial coverage.

2.3. Santa Ana winds chronology

Historical days when a Santa Ana wind event is known to have occurred were isolated using a previously published chronology [11]. Santa Ana days in this dataset were identified based on synoptic meteorological criteria including a strong northeastward pressure gradient across the southwestern part of California and high cold-air advection above the Southern California deserts [11]. While there are several published reconstructions of Santa Ana days, we chose this one because of its accessibility, ease of use, and temporal depth (reconstructed Santa Ana days from 1950-2010).

2.4. Fire occurrence data

The fire occurrence database (FOD) is a comprehensive collection of wildfire attributes for the United States constructed from national, state, and private fire records [44]. From this database, we extracted final fire size, ignition location, and discovery date for all large fires ($\geq$ 40.47 ha) with a recorded ignition location within the Pyrome 34 boundary (Figure 1). The database lists all fires from 1992-2015, but we included only fires that ignited between 1996-2010 in order to analyze in conjunction with the RAWS and Santa Ana wind days datasets.

2.5. Wind variability during Santa Ana and non-Santa Ana fires

We merged the days of recorded historic Santa Ana wind days with both the RAWS wind records and the FOD to construct four fire-wind categories: a) Santa Ana days with a large-fire ignition (LF-SA); b) Santa Ana days without a large-fire ignition (noLF-SA); c) non-Santa Ana days with a large-fire ignition (LF-noSA); and d) non-Santa Ana days without a large-fire ignition (noLF-noSA). We divided our data into these four categories because mounting evidence shows that environmental factors responsible for fire vary significantly between Santa Ana and non-Santa Ana fires [7, 22, 23]. Merging the three datasets (RAWS, FOD, and Santa Ana winds chronology) results in 24 h of wind vectors each day for each station between 1996-2010. We acknowledge that merging only the fire ignition date with the Santa Ana chronology may not fully capture winds over the lifetime of a fire, and future work would benefit from assessing the temporal patterns of wind heterogeneity in relation to hourly or daily fire spread.

For each of the four fire-wind categories, we summarized between- and within-station wind directions and speeds for all 30 RAWS. Between-station wind summaries are presented statistically as means of all hourly RAWS records during each fire-wind category and visually via wind roses; within-station wind summaries are presented as means and wind roses at each individual station.

Two (one for wind speed and one for wind direction) non-parametric Kruskal-Wallis tests were implemented in R to confirm that there were statistically significant differences (p-value < 0.01) between wind speeds and wind directions across the four fire-wind groups. Neither wind speed nor direction are normally distributed and variances are unequal across the four fire-wind categories, precluding use of traditional parametric methods like ANOVA. A pairwise posthoc Mann-Whitney-Wilcoxon test was also conducted to determine the magnitude of difference between each of the fire-wind category pairs (Table S2.4).

2.6. Clustering of weather stations into groups with similar wind vectors

To explore whether the RAWS clustered into groups of similar wind vectors, we implemented an agglomerative hierarchical clustering algorithm. Hierarchical clustering aims to group stations with similar wind characteristics and does not attempt to explain underlying drivers of station heterogeneity. We used the "hclust" function in R on both wind direction and speed for each of the four fire-wind categories [45]. For direction and speed separately, the function first constructs dissimilarity matrices across the 30 RAWS, which reports the similarity between each pairing of weather stations. Agglomerative clustering works by assigning each station's wind stream its own individual cluster; the algorithm then continues to iteratively join the next closest station in progressive stages until just one cluster with all 30 weather stations remains. Distances between groups were calculated using Ward's minimum variance method, which allocates clusters by minimizing the within-cluster variance [46].

We present the results of each hierarchical clustering as dendrograms to visualize the station groupings. To calculate the potential optimum number of clusters we applied the "elbow" method, which attempts to identify the number of clusters k that minimize total within-cluster sum of squares. The hierarchical clustering algorithm was repeated for k 1–10 clusters, and the total within-cluster sum of squares calculated for each k. We plotted total within-cluster sum of squares against the number of clusters; the point at which the curve bends, i.e. the "elbow," is considered an appropriate number of clusters, although determining this number is a visual and subjective distinction [47]. We then compared the "elbow" curves visually with the dendrograms to interpret the final clustering.

2.7. Station siting heterogeneity

We additionally explored variations in basic geographic characteristics of the 30 RAWS: slope position (upper = top 1/3 of slope; middle = middle 1/3 of slope; lower = lower 1/3 of slope), elevation (m), aspect (flat, N, NE, E, SE, S, SW, W, NW), and climate class (Arid/Semiarid, Subhumid (summer rain deficient), Subhumid (year-round rainfall), and Wet). Each of these station traits are recorded by the managers of the individual RAWS and included in station metadata [48]; we extracted the variables directly from the metadata without additional post-processing.

Additionally, an extended description of our statistical methodology is provided in AppendixS1.

3. Results

3.1. Variability between wind-fire categories and among individual stations

LF-SA days are the rarest of the four fire-wind categories; only 1% of all hours analyzed fall into this category (Table 1). NoLF-noSA are most common, with approximately 85% of all hours analyzed. Santa Ana days in general are rare, and only comprise 8% of hours. Most days do not experience an ignition that leads to a large fire (Table 1).

A Kruskal-Wallis test confirmed that wind vectors are significantly different (p-value < 0.01 for both wind speed and direction) across each of the four fire-wind categories (Table S2.3). All pairwise Mann-Whitney-Wilcoxon comparisons are also significantly different (p-value < 0.01). The statistical differences in the speed and direction distributions of each of the fire-wind categories (i.e. the location parameters of each Mann-Whitney-Wilcoxon test, which estimates the median of all differences between samples from each category pair) support statistical significance for all pairwise comparisons except between LF-noSA and noLF-noSA (Table S2.4). The large sample size of noLF-noSA (>3,000,000 h) likely leads to a small p-value, even though the location parameter (-3.15e-07) is too small to assume a practical difference based on statistical significance alone.

As expected, average wind speeds on both LF-SA and noLF-SA days are greater than non-Santa Ana days (Table 1, Figure 2). Also as expected, wind directions are largely confined to the northeast quadrant (330 - 105°) on Santa Ana days, while on non-Santa Ana days they span all directions. Interestingly, we observe a statistically significant difference in the average wind speed between days with large fire ignitions and without (Table 1). During Santa Ana days, wind speeds are higher on days with a large fire (LF-SA) by an average of 1.23 m/s (31%) than days with no large fires (noLF-SA). When there is no Santa Ana wind, the average wind speed is actually higher when there is no large fire ignition (noLF-noSA) than when there is a large fire ignition (LF-noSA) (Table 1).

Wind directions are closer to true northeast (i.e. 45°) on LF-SA days than the other three fire-wind categories, including noLF-SA days (6.49° from true northeast on large-fire ignition days versus 11.46° from true northeast on non-large fire ignition days). The difference in the average wind directions between Santa Ana days and non-Santa Ana days is nearly a complete directional reversal (38.51° and 33.54° on LF-SA and noLF-SA days versus 234.21° and 236.92° on LF-noSA and noLF-noSA days), although during non-Santa Ana days the directions are well-populated throughout the entire compass rose (Figure 2).

Across Southern California, heterogeneity appears in the wind vectors recorded by individual weather stations (Figure 3). Within and across individual stations, winds also vary during LF-SA days (Figure S3.1), noLF-SA days (Figure S3.2), LF-noSA days (Figure S3.3), and noLF-noSA days (Figure S3.4). Average wind speeds on LF-SA days can exceed 10 m/s (e.g. Chilao, Fremont Canyon, Warm Springs RAWS), or be less than 2 m/s (e.g. Mt. Laguna, Oak Grove, Rose Valley RAWS), a difference of over 8 m/s within the same fire-wind category (Table S2.1). At all stations, wind speeds are consistently higher on LF-SA days than during any other fire-wind category; at most stations, winds also show a pronounced northeasterly directional concentration. Winds are consistently slower on non-Santa Ana days, and there are less obvious within-station differences on LF-noSA days versus noLF-noSA days. Wind directions on non-Santa Ana days are typically characterized by a bi-modal directional distribution at many stations, a pattern that is not evident by aggregating all stations together (Figures S3.3 and S3.4 versus Figure 2). In contrast, on Santa Ana days winds tend to have a unimodal directional peak, with the exception of the bimodal wind directions at Montecito (Figure 3, Figure S3.1).

3.2. RAWS cluster into different groups for each fire-wind category

We find some evidence that RAWS cluster into reasonably distinct categories on LF-SA (Figure 4), noLF-SA (Figure 5), LF-noSA (Figure 6), and noLF-noSA (Figure 7) days. Using the "elbow" method, we determine that 2–3 clusters are most likely to reduce total within-group sum of squares. Visual examination of the station data revealed that three clusters, rather than two, are probably the most practical grouping in most cases; however, two clusters seem most reasonable for LF-noSA and noLF-noSA fire clusters (Figure 6 and Figure 7).

Stations do not cluster into identical groups across all categories. On LF-SA days, stations group into one of three distinct wind direction categories: 1) strong northeast peak with little variability around the compass rose; 2) strong north/northwest peak with little variability around the compass rose; and 3) high variability around the compass rose with a small northeast peak (Figure 4). These same three directional categories emerge on noLF-SA days as well, but the stations re-group slightly (Figure 5). Wind speeds on LF-SA days clearly separate into high-speed (approximately >6.70 m/s), medium-speed (<6.70 m/s and >4.47 m/s) and low-speed (<4.47 m/s) clusters. These same three wind speed categories emerge on noLF-SA days, but overall wind speeds are generally lower: high-speed (>5.36 m/s), medium-speed (<5.36 m/s and >2.24 m/s), and low-speed (<2.25 m/s).

On non-Santa Ana days, station clusters are less obvious, and all stations exhibit a bimodal wind direction distribution. Generally, stations group into one of two wind direction categories on LF-noSA days: 1) bimodal distribution with prominent northeast/southwest or north/south peaks; or 2) bimodal distribution with a prominent northeast peak (Figure 6). These two groupings also emerge on noLF-noSA days. The "elbow" curve suggests that two or three clusters may be appropriate for non-Santa Ana wind speeds, but examination of the clusters and underlying data suggest that two clusters are most appropriate. On both LF-noSA and noLF-noSA days, stations group into clusters of relatively high (>1.56 m/s) and low (<1.56 m/s), but no clear "medium-speed" category emerges.

3.3. Station siting

Examination of geographic variables reported in the station metadata also revealed interesting patterns of station siting (Figure S3.5, Table S2.2). Stations were fairly equally distributed across elevation and slope position, but the distribution was skewed in terms of aspect and climate class. Flat (n = 11) and south-facing (n = 2, n = 7, and n = 6 for SE, S, and SE, respectively) aspects were overrepresented, accounting for 26 of the 30 stations combined. This placement is by design, since RAWS are sited to represent extreme fire conditions; south slopes in the northern hemisphere tend to have the most extreme fire conditions since their higher insolation contributes to low fuel

Table 1. Means of hourly wind vectors observed at 30 Southern California RAWS from 1996-2010 during Santa Ana days (fires and non-fires) and non-Santa Ana days (fires and non-fires). 95% confidence intervals are in brackets.

	Wind direction (°)	Wind speed (m/s)	Proportion of Hours
LF-SA	38.51 [38.48, 38.53]	5.19 [5.14, 5.23]	0.01
noLF-SA	33.54 [33.53, 33.55]	3.96 [3.94, 3.97]	0.07
LF-noSA	234.21 [234.15, 234.27]	2.30 [2.29, 2.31]	0.07
noLF-noSA	236.94 [236.92, 236.96]	2.38 [2.38, 2.38]	0.85

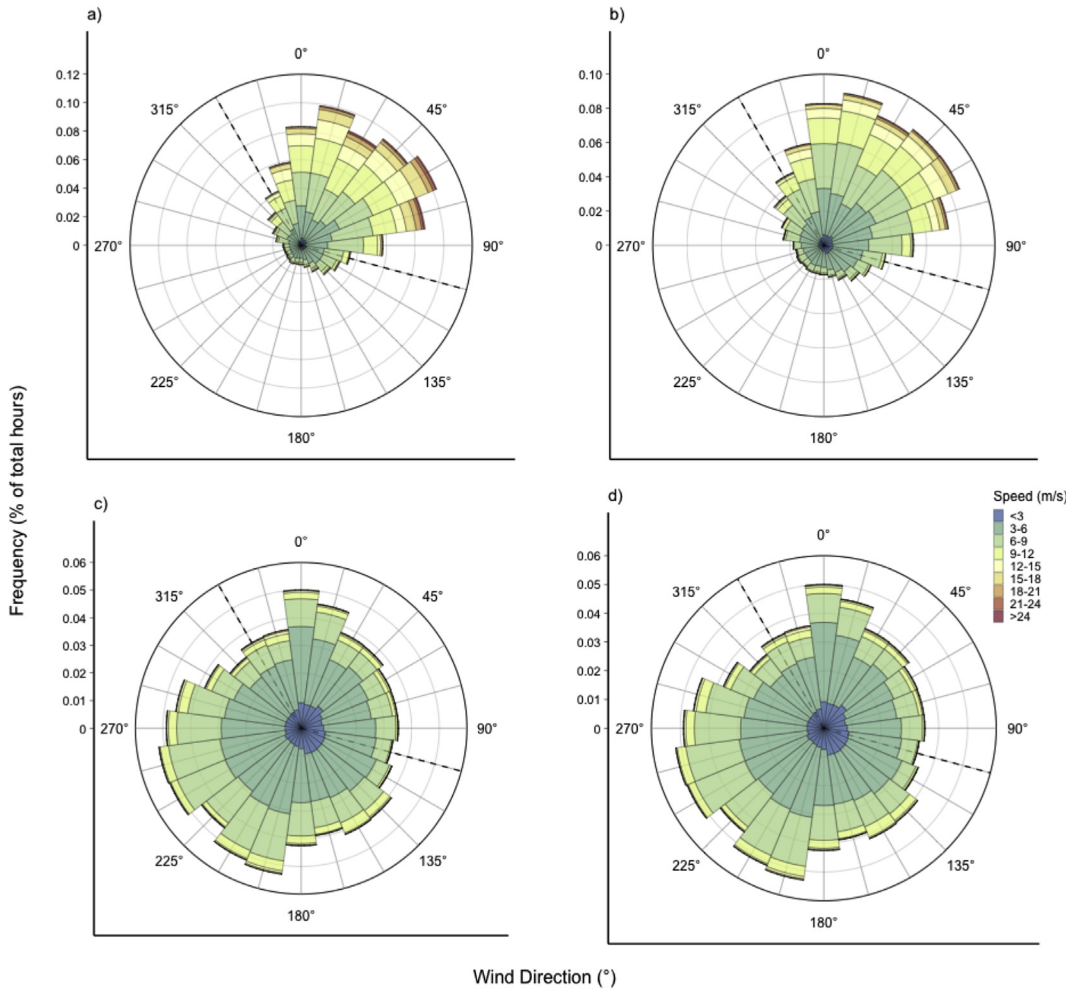

Figure 2. Distribution of hourly wind directions and wind speeds aggregated across all 30 RAWS from 1996-2010 on a) Santa Ana days with a large fire ignition (LF-SA); b) Santa Ana days without a large fire ignition (noLF-SA); c) non-Santa Ana days with a large fire ignition (LF-noSA); and d) non-Santa Ana days without a large fire ignition (noLF-noSA). Height of the bars on the wind rose represents the percentage (y-axis) of total hours that are within each 15° directional bin. Colors of the bars indicate the percentage of total hours within each wind speed bin.

moisture. Most stations are located in a subhumid (summer rainfall deficient) climate class (n = 27), which corresponds with the dominant climate across most of this region. However, three stations are in arid/semiarid climates, which should be noted when considering these RAWS for modeling fire in a predominantly subhumid climate region.

Of the 4 siting variables, only slope position is consistently related to wind speeds, with "upper" slope stations typically reporting higher wind speeds (Figure 3, Table 2). However, this could also be related to unequal siting of "upper" slope stations across the landscape; RAWS in the mountains to the north and west of Los Angeles, where Santa Ana winds are traditionally strongest, tend to be sited on upper slopes while also recording high-speed northeast or northwest winds during LF-SA days. In contrast, the southernmost RAWS tend to exhibit lower-speed eastern winds on LF-SA days, with no stations located on upper slopes (Figure 3).

4. Discussion

By combining observed hourly wind vectors from meteorological stations with fire occurrence data, we show that winds vary geographically across southern California not just during Santa Ana and non-Santa Ana days, but also between days with large fire ignitions and those without. This result has key implications for understanding regional wind variability as well as improving the use of observational meteorological datasets in large-fire simulation modeling. Our findings help summarize some aspects of wind-fire heterogeneity in southern California, identify areas where fire simulation models may be limited, and contribute to improving practical implementations of the results of fire risk simulations.

Across southern California, we notice several emergent wind characteristics: a) wind directions are closer to northeast (i.e. 45°) during LF-SA days than any other wind-fire category; b) mean wind speeds are 31% higher on LF-SA days (5.19 m/s) than noLF-SA days (3.96 m/s); c) mean wind direction differs by nearly 180° on Santa Ana days (38.51° and 33.54° on LF-SA and no-LF-SA days, respectively) versus non-Santa Ana days (234.21° and 236.92 on LF-noSA and noLF-noSA days, respectively); d) on non-Santa Ana days, many RAWS have a bimodal wind distribution; and e) there is greater variability around the compass rose for winds during non-Santa Ana days than Santa Ana days. Although variability exists between stations on Santa Ana days, it is more concentrated in a northeasterly direction for all RAWS than on non-Santa Ana days.

These characteristics of observed winds allow us to corroborate and add detail to results reported by others. In an analysis of fire spread data for 30 Santa Ana and 26 non-Santa Ana fires from 2002-2009, Santa Ana fires were spread primarily by northeast and east winds, but with directional variability around the entire compass rose [7]. In contrast, spread of non-Santa Ana fires is concentrated in an east-southeast

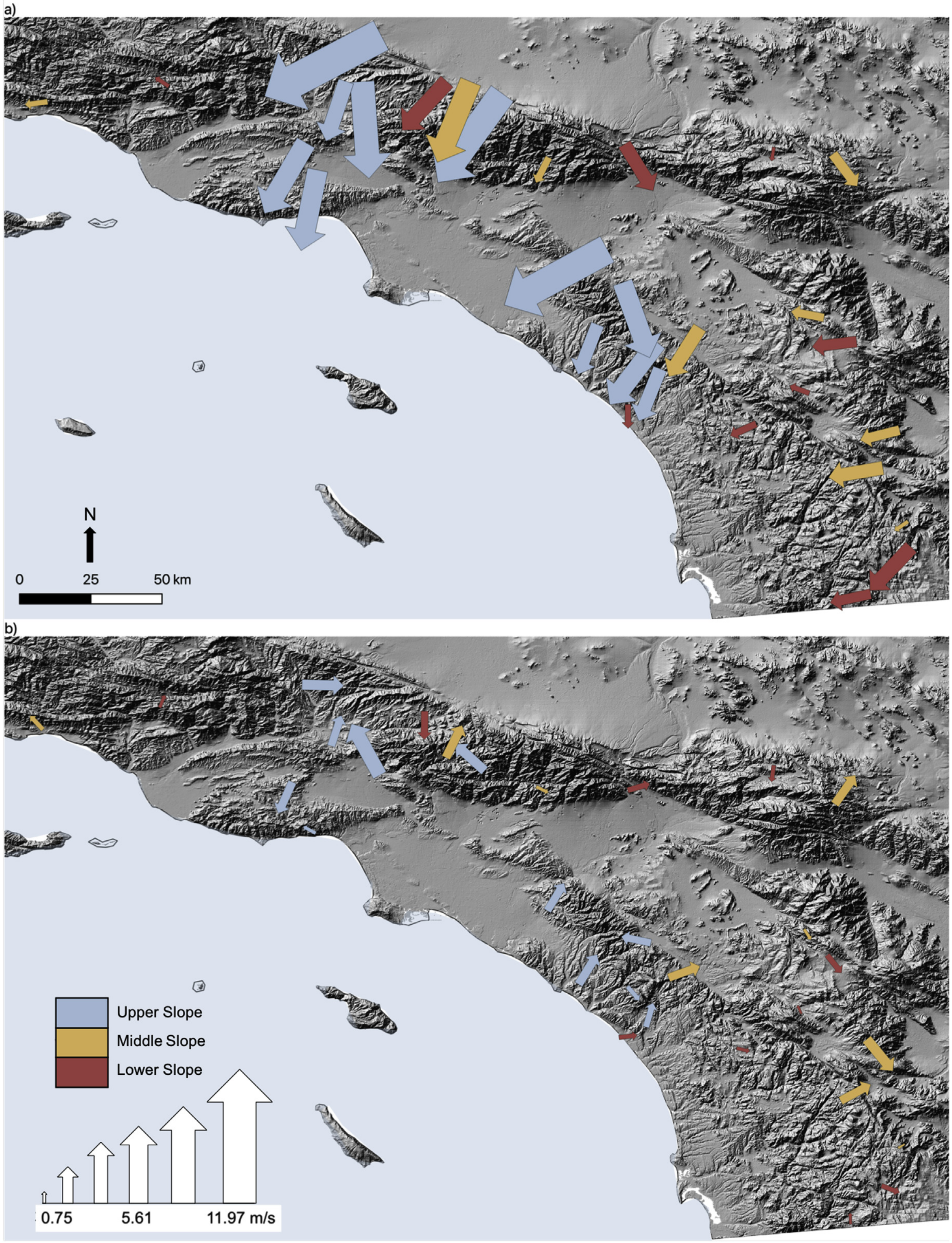

Figure 3. Map of the average wind direction and speed recorded at each RAWS during a) Santa Ana days with a large fire ignition (LF-SA); and b) non-Santa Ana days with a large fire ignition (LF-noSA). Arrow direction indicates the mean direction of wind flow using meteorological convention (i.e. an arrow pointing south represents a north wind). Arrow size indicates mean wind speed on a gradient from a minimum speed of 0.75–11.97 m/s. Arrow colors indicate ordinal slope position of the station site as recorded in RAWS metadata: upper = highest 1/3 of slope; middle = middle 1/3 of slope; lower = lowest 1/3 of slope.

direction, probably due to the prevailing onshore wind patterns. Others also found that wind direction on the first day of Santa Ana fires averaged north at 353°, while on days without Santa Ana conditions, winds shifted to an average of 234° [19]. Both of these findings are corroborated by the patterns we see among all RAWS.

Many studies have highlighted the dominant role of Santa Ana winds in the spread of large fires in Southern California [20, 49, 50]. Indeed, wind speed is a reliable indicator of fire spread in chaparral vegetation, especially on the first few days following ignition [19, 51, 52]. Wind speeds on LF-SA days are higher than during any other wind-fire

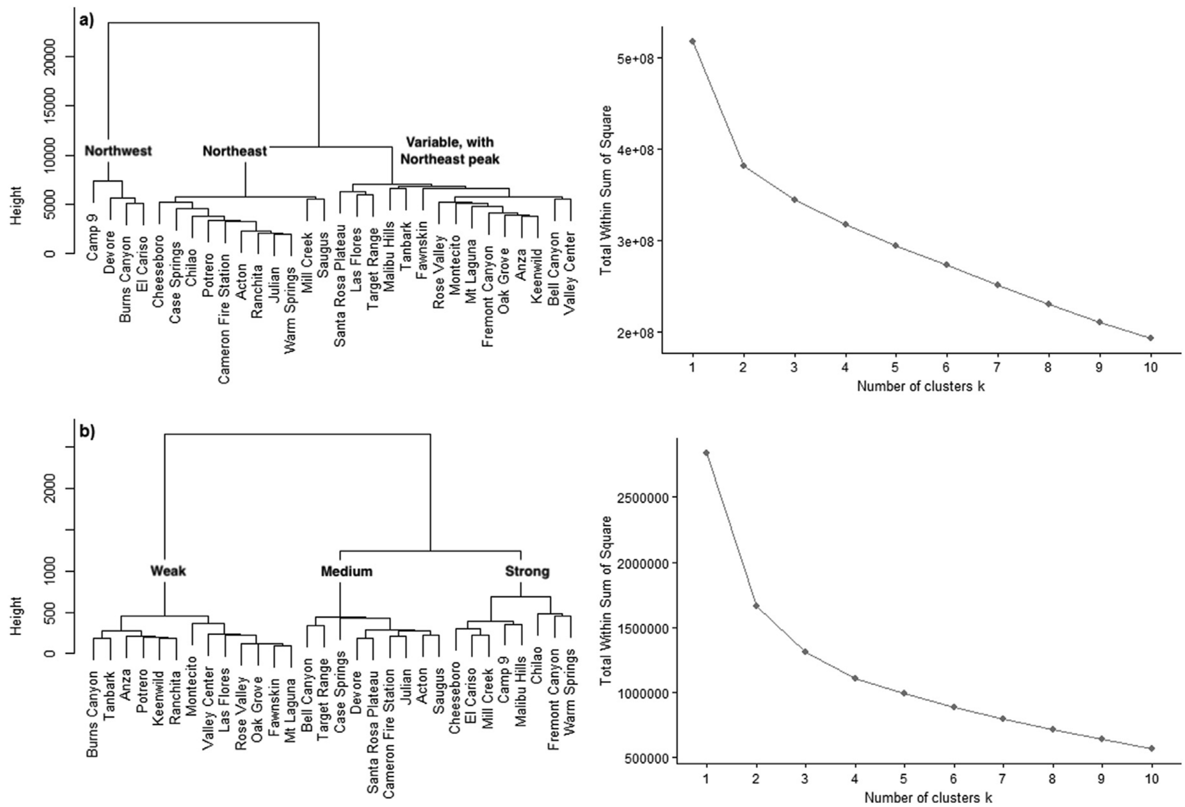

Figure 4. Agglomerative hierarchical clustering results for a) wind direction and b) wind speed during LF-SA days. In the left panel, total within-clusters sum of squares is plotted for 1–10 cluster groups. Labels on the dendrograms indicate the clustering groups identified from the clustering analysis and assessment of the wind roses. The right panel shows the total within-cluster sum of squares plotted against the number of clusters. The number of clusters at which the curve bends, or the "elbow," is considered an acceptable number of clusters.

category, including noLF-SA days (Table 1, Table S2.1), confirming that stronger winds are more likely to be associated with large fires. Many of the stations with the highest Santa Ana wind speeds (e.g. Warm Springs, Camp 9, Cheeseboro) are similarly placed in the mountainous areas north and west of Los Angeles where many large fires occur (Figure 1). Winds on LF-noSA days are actually weaker than on noLF-noSA days, showing that wind strength in the early growth of the fire is a less important factor in the initiation of non-Santa Ana fires, possibly exceeded by drought.

Santa Ana winds are apparent at nearly all RAWS, and the great majority of the wind directions recorded during Santa Anas are from the northeast quadrant (Figures 4, 5, 6, and 7). Where stations are partially constrained by directional topographic features like mountains and valleys, prevailing wind directions during Santa Ana wind events may lie outside the conventional northeast quadrant (e.g., Rose Valley and Valley Center, Figures 4 and 5). This suggests that although most stations manage to capture Santa Ana conditions, selecting one or two representative stations requires some careful analysis. On non-Santa Ana days, the between-station variability is high, suggesting that local wind conditions are primarily constrained by local topography instead of larger synoptic patterns.

On LF-SA and noLF-SA days, RAWS separate into three fairly distinct directional categories: strong northeast peak with low variability, strong north/northwest peak with low variability, and moderate northeast peak with variability around the compass rose (Figure 4, Figure 5). Likewise, RAWS separate into three clusters of high, medium, and low wind speeds. These clustering groups in part reflect some of the underlying geography of southern California (Table S2.1).

During Santa Ana events, winds have been shown to channel through major gaps in the southern California mountains, including the Soledad, Cajon, and San Gorgonio/Banning Passes [25], down the Santa Clarita River Valley [17], and around the Laguna Mountains [26]. These areas are also associated with heightened fire danger because of the increased local strength of Santa Ana winds [26]. In contrast, areas including the lower-lying basin surrounding Los Angeles bound by the foothills of the San Gabriel, Santa Ana, and San Bernardino Mountains and the coastal hills to the northwest between Los Angeles and Santa Barbara, are typically characterized by weaker and more variable Santa Ana winds and associated fire danger [13, 18, 24, 26, 53]. In the San Gabriel Mountains, stations have recorded changing wind characteristics over the course of the same series of Santa Ana days [25]. High-elevation stations near the Soledad Pass consistently recorded high-speed Santa Ana winds from the northeast, while winds at stations located in canyons and on the eastern slope of the mountains generally recorded variable wind directions and lower wind speeds during all but the most intense Santa Ana conditions. The Tanbark RAWS station that we analyzed (near the Tanbark Flats station analyzed in [25]), exemplifies the characteristic variability in winds that occur on the eastern slope of the mountains.

In our analysis of the RAWS wind vectors, some of these same general geographic patterns emerge (Figure 3, Table S2.1). Most of the RAWS that experience strong Santa Ana wind characteristics are located near high mountain passes, in mountain ranges, and in areas that are otherwise associated with high Santa Ana wind strength and frequency (Figure 1, Figure 3). Winds at stations on the eastern slopes of the mountains, near the coast, or in valleys, basins, and canyons tend to be more variable with semi-frequent occurrence of opposing wind directions as well as generally lower wind speeds.

For example, near the Soledad Pass, RAWS at Warm Springs, Saugus, and Acton all exhibit strong northeast winds and high wind speeds during LF-SA and noLF-SA days (Figures 3, 4, and 5). Likewise, RAWS on the eastern periphery of Soledad Pass, including Chilao and Mill Creek, record high-speed winds with a strong northeast gradient. Camp 9 is an exception that falls into the north/northwest directional cluster, but still records wind speeds in the high-speed category. However, Camp 9 sits near the outlet of a northwest/southeast oriented valley, which is likely

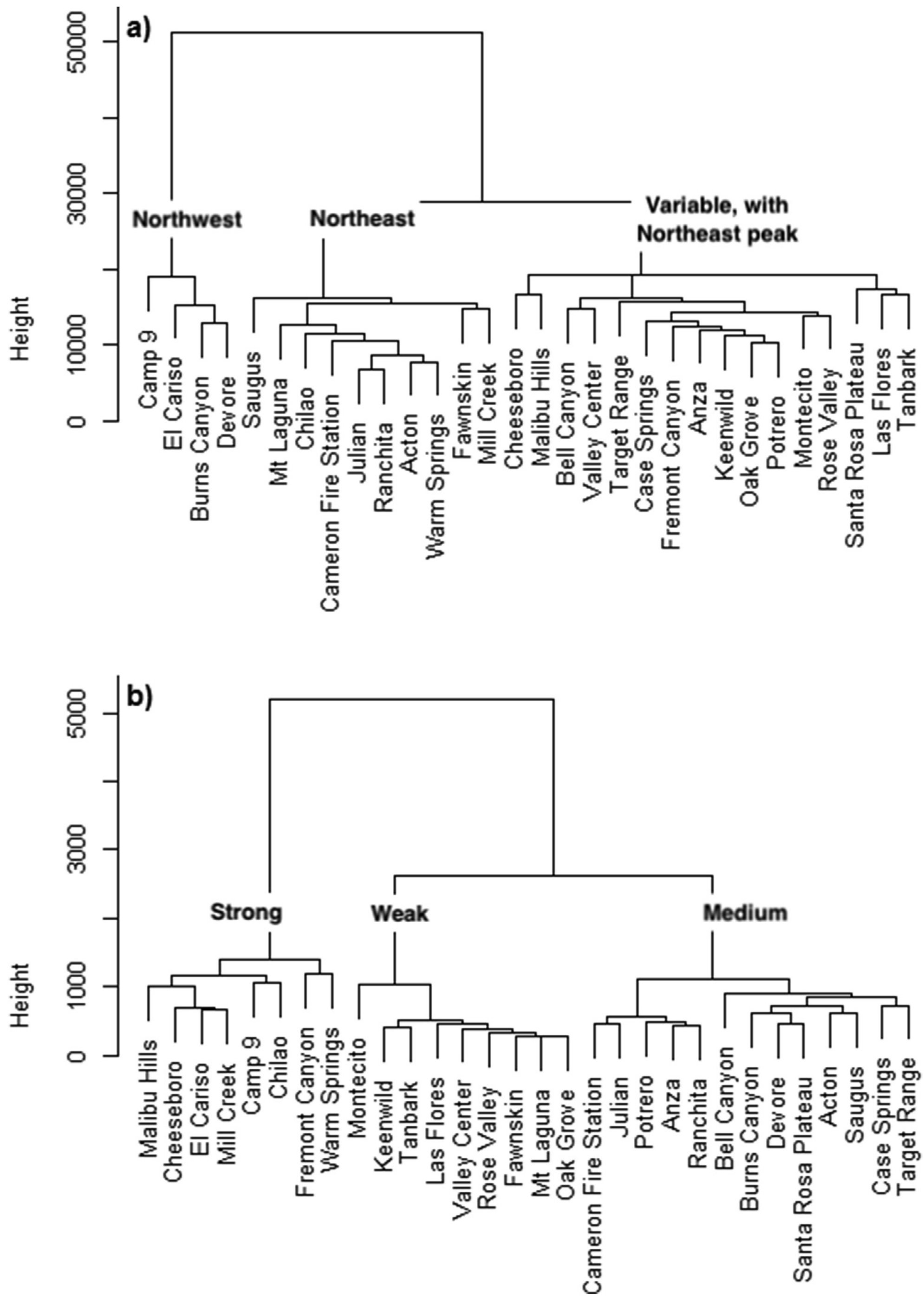

Figure 5. Agglomerative hierarchical clustering results for a) wind direction and b) wind speed during noLF-SA days. Labels on the dendrograms indicate the clustering groups. Within-cluster sum of square "elbow" plots are similar to those shown in Figure 4, and omitted here.

responsible for its observed directional pattern. Near the outlet of the Soledad Pass in the Santa Clarita River Valley is the Cheeseboro RAWS, which falls into both the strong northeast directional gradient and high-speed categories. Devore is the sole RAWS located directly in the Cajon Pass; the strong north/northwest gradient shown here is consistent with the orientation of the pass and the wind directions known to come through this pass [26]. We did not analyze any RAWS directly within the San Gorgonio/Banning Pass. However, the two closest RAWS, Keenwild and Anza, have an easterly directional peak with variability around the compass rose, consistent with the direction that Santa Ana winds would be expected to funnel through the San Gorgonio Pass. In the southern Laguna Mountains, we observed a predominance of the classic northeasterly Santa Ana winds at the Potrero, Ranchita, Cameron Fire, Julian, and Mt. Laguna RAWS, although wind speeds tend to classify in the medium- and low-speed categories at these stations (Figures 4 and 5, Figure S3.1, and Figure S3.2).

Outside of the major mountain ranges in the lower-lying areas in and around the Los Angeles basin, winds at the Oak Grove, Valley Center, Santa Rosa Plateau, and Tanbark RAWS are more variable and weaker. Although northeast winds are still predominant, the higher variability is consistent with supporting evidence that Santa Ana wind-driven fire appears to be less intense and/or more variable in these physical environments [24, 25, 26]. In the northwest part of our study area, the Montecito and Rose Valley stations also exhibit directional variability

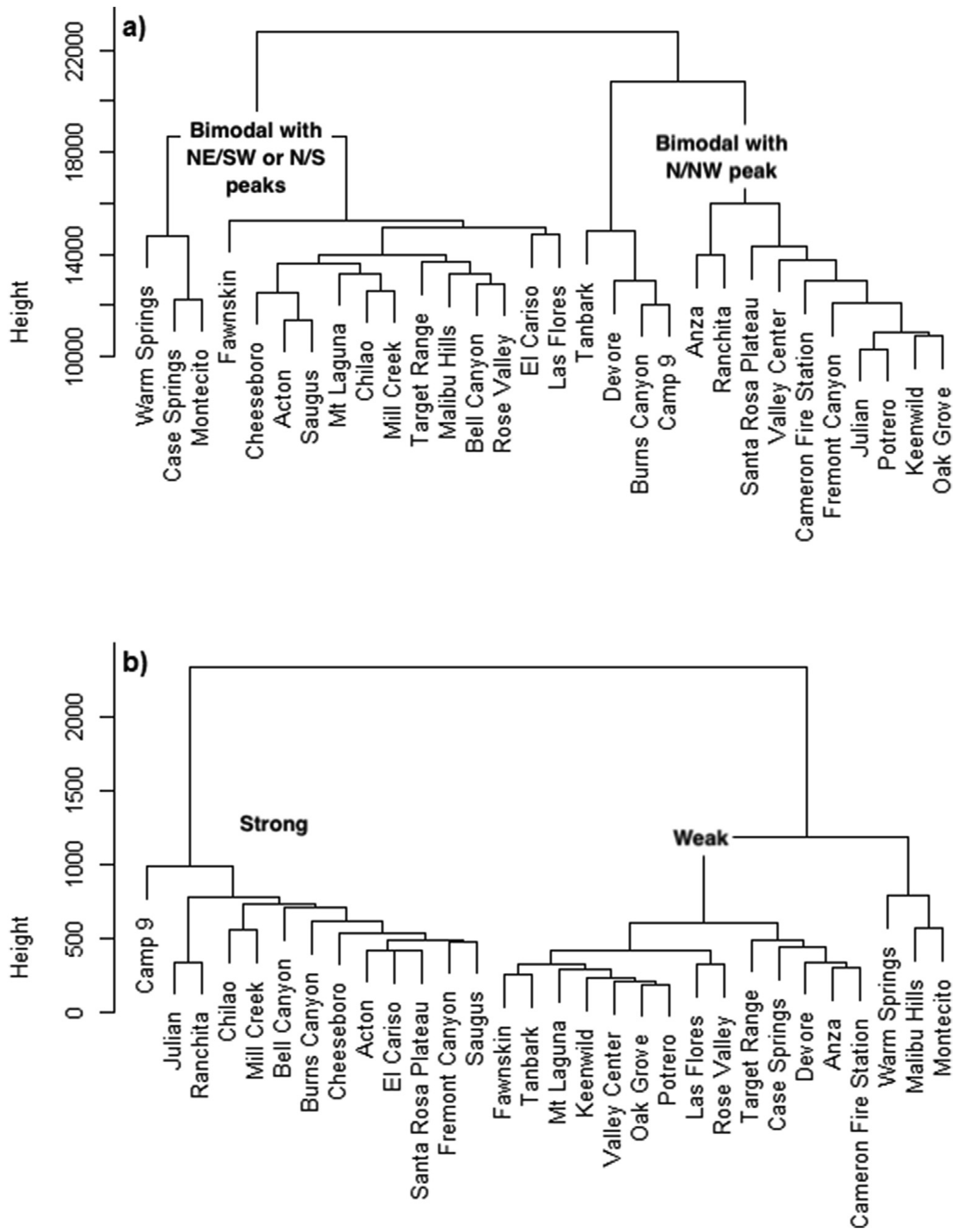

Figure 6. Agglomerative hierarchical clustering results for a) wind direction and b) wind speed during LF-noSA days. Labels on the dendrograms indicate the clustering groups. Within-cluster sum of square "elbow" plots are similar to those shown in Figure 4, and omitted here.

and low wind speeds. In this area, Santa Ana winds may abate or coexist with counteracting on-shore breezes [18, 23, 53]. Due to its coastal location, Montecito appears to be the only RAWS that exhibits a clear bimodal distribution of wind directions during Santa Ana days, with an "on-shore" breeze component from the southwest occurring on approximately 50% of its recorded hours (Figure S3.1).

On non-Santa Ana days, wind vectors exhibit a much less obvious association with the regional-scale topography of southern California, but are influenced highly by local topography. More so than on Santa Ana days, many stations exhibit some degree of bimodality with opposing directional peaks that are consistent with their geographic location within a valley, ridge top, or mountain slope (Table S2.1, Figure S3.7). Despite clearly different wind concentrations overall, there are some notable connections between stations on non-Santa Ana and Santa Ana days. For example, stations that exhibited a strong north and northwest pattern on Santa Ana days are prone to exhibit a clear bimodal north-south pattern on non-Santa Ana days. The most common directional cluster has a fairly consistent northeast peak accompanied by a highly variable second mode; these patterns are directly related to the geographic location of the stations within directionally oriented canyons or ridges (Table S2.1). Stations cluster primarily into "high" and "low" speed classes on non-Santa Ana days, with no distinct "medium" cluster (Figures 6 and 7). At most stations, winds are weaker on non-Santa Ana days than on Santa Ana days and are generally recorded across all directions that are geographically feasible given the station's topographic conditions.

The high degree of spatial heterogeneity among key model variables is a perplexing issue at the forefront of fire-climate model development [9]. Wind is one of the primary meteorological variables used to model fire spread [34, 52, 54]. The wind heterogeneity observed across 30 weather stations helps interpret the conditions each station experiences. In a post-analysis survey of station location characteristics, we applied a

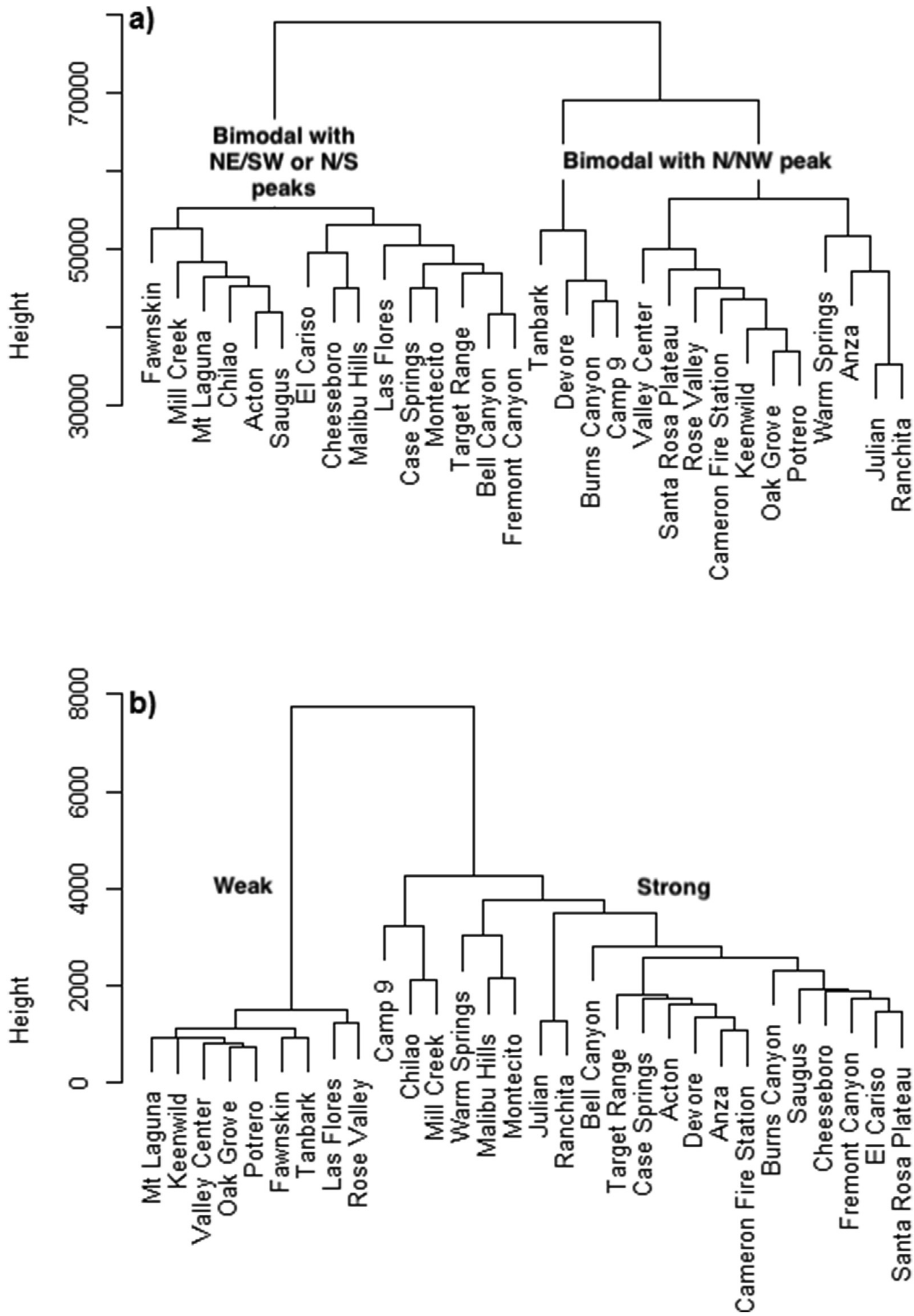

Figure 7. Agglomerative hierarchical clustering results for a) wind direction and b) wind speed during noLF-noSA days. Labels on the dendrograms indicate the clustering groups. Within-cluster sum of square "elbow" plots are similar to those shown in Figure 4, and omitted here.

subjective decision-making process to determine which stations are likely to be suitable for fire simulation modeling. In general, stations should be positioned where they reliably capture the synoptic wind patterns at any point in time with the capability of recording winds from all directions [34, 42]. These criteria relegate stations constrained by canyons, narrow valleys, or mountain slopes in favor of ridge top and broad valleys sites.

We initially identified five RAWS that would meet the geographic criteria for use in fire simulation models: Case Springs, Chilao, Keenwild,

Table 2. Mean wind speed (m/s) for upper, middle, and lower slope positions on LF-SA and LF-noSA days.

	Upper	Middle	Lower
LF-SA	4.89	2.41	2.10
LF-noSA	1.87	1.48	1.00

Saugus, and Valley Center (Table S2.2). Each of these stations is located on a ridge top or in a broad valley with minimal potential for topographic blocking or channeling of winds. These choices are further supported by their wind rose distributions; on Santa Ana and non-Santa Ana days, each of the five stations records winds from all directions, although they each have a clear concentration of northerly winds on Santa Ana days (Table S2.1, Figures S3.1, S3.5, S3.6, S3.7).

The other 25 RAWS provide a template for how wind vectors vary across a diverse landscape like Southern California. Winds at these stations reflect in part the unique topographic characteristics of their locations (Table S2.2). Some geographic environments are underrepresented (Figure S3.5). For example, there are few stations with NW or NE/E aspects; and, although there is a fair overall distribution of upper, middle, and lower slope positions, these do not appear to be equally distributed geographically (Figure 3, Figure S3.5). Stations situated high on a slope (not necessarily higher elevation, but relative to their respective slopes) tend to have higher average wind speeds (Table 2). However, most of the upper slope RAWS are also located in the region of traditionally highest Santa Ana wind activity [25, 26, 37].

As computational capabilities increase, it may become possible to use multiple stations to reflect spatial variability in wind in fire simulations, including the relationship between Santa Ana winds and fire. Furthermore, modeling fire spread using data from a station that doesn't capture synoptic conditions can drastically alter the number, size, and total burned area of simulated fires [30]. Our station-level analysis provides a template that fire modelers can use to select observational data appropriate to their purpose. For example, to model fires across all of Southern California, one may select from the five stations we suggested as good candidates to record synoptic-scale winds. Another approach could be to include wind data from several weather stations — for example, using data from stations representing a variety of wind vectors, slope positions, and aspects — to assess the impact of geographic wind variability on fire spread. Alternatively, if simulating spread of a specific fire (e.g., an ignition in the Cajon Pass or the Laguna Mountains) is of interest, our station-level wind summaries can provide users with critical information regarding the direction and speed at which a fire is likely to spread.

5. Conclusion

The natural fire regime of southern California is an integral component of the human and ecological landscape, and winds are one of the key meteorological variables (along with fuel moisture, vegetation patchiness, ignition source, and others) responsible for the rate and direction of fire spread. In southern California, surface winds are complicated by the Santa Ana wind phenomenon, as well as the complex local topography. Simulating large fire risk in this region requires an appreciation of this complexity.

Here, we have presented a detailed summary of wind variability at 30 meteorological stations on the southern California landscape during Santa Ana and non-Santa Ana days with and without a large fire ignition. To date, no published research has taken advantage of the widespread modern observational data networks that are currently available to analyze wind variability during fire ignitions. Our work suggests explanations for wind heterogeneity across southern California and provides a template for appropriate station selection as inputs to wind and fire simulation models. We intend for this work to inform fire simulation modeling efforts in the future, as well as instigate additional research of the local meteorological factors that are responsible for fire spread in this region.

Declarations

Author contribution statement

Alex W Dye: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

John B. Kim, Karin L. Riley: Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Funding statement

Alex W. Dye is a fellow with the Oak Ridge Institute for Science and Education (ORISE), supported by the USDA Forest Service Western Wildland Environmental Threat Assessment Center (WWETAC). John B. Kim is supported in part by WWETAC.

Competing interest statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2020.e04159.

References

[1] S.A. Mensing, J. Michaelsen, R. Byrne, A 560-year record of Santa Ana fires reconstructed from charcoal deposited in the Santa Barbara Basin, California, Quat. Res. 51 (1999) 295–305.
[2] R.A. Minnich, Fire mosaics in southern California and northern Baja California, Science 219 (1983) 1287–1294 (80).
[3] A.D. Syphard, J.E. Keeley, A.B. Massada, T.J. Brennan, V.C. Radeloff, Housing arrangement and location determine the likelihood of housing loss due to wildfire, PloS One 7 (2012).
[4] R.J. Delfino, S. Brummel, J. Wu, H. Stern, B. Ostro, M. Lipsett, A. Winer, D.H. Street, L. Zhang, T. Tjoa, D.L. Gillen, The relationship of respiratory and cardiovascular hospital admissions to the Southern California wildfires of 2003, Occup. Environ. Med. 6 (2009) 790–795.
[5] J.C. Liu, L.J. Mickley, M.P. Sulprizio, X. Yue, R. Peng, F. Dominici, M.L. Bell, Future respiratory hospital admissions from wildfire smoke under climate change in the Western US, Environ. Res. Lett. 11 (2016) 124018.
[6] D.E. Calkin, K.M. Gebert, J.G. Jones, R. Mountain, Forest Service large fire area burned and suppression expenditure trends, 1970-2002 103 (2005) 179–183.
[7] Y. Jin, M.L. Goulden, N. Faivre, S. Veraverbeke, F. Sun, A. Hall, M.S. Hand, S. Hook, J.T. Randerson, Identification of two distinct fire regimes in Southern California: implications for economic impact and future change, Environ. Res. Lett. 10 (2015), 094005.
[8] A.L. Westerling, B.P. Bryant, Climate change and wildfire in California, Climatic Change 87 (2008) 231–249.
[9] J. Keeley, A. Syphard, Climate change and future fire regimes: examples from California, Geosciences 6 (2016) 37.
[10] J. Guzman-Morales, A. Gershunov, Climate change suppresses Santa Ana winds of Southern California and sharpens their seasonality, Geophys. Res. Lett. (2019).
[11] J.T. Abatzoglou, R. Barbero, N.J. Nauslar, Diagnosing Santa Ana winds in southern California with synoptic-scale Analysis, Weather Forecast. 28 (2013) 704–710.
[12] M.N. Raphael, The Santa Ana winds of California, Earth Interact. 7 (2003) 1–13.
[13] J. Guzman-Morales, A. Gershunov, J. Theiss, H. Li, D. Cayan, Santa Ana Winds of Southern California: their climatology, extremes, and behavior spanning six and a half decades, Geophys. Res. Lett. (2016).
[14] S. Conil, A. Hall, Local regimes of atmospheric variability: a case study of Southern California, J. Clim. 19 (2006) 4308–4325.
[15] M. Hughes, A. Hall, Local and synoptic mechanisms causing Southern California's Santa Ana winds, Clim. Dynam. 34 (2010) 847–857.
[16] Y. Cao, R.G. Fovell, Downslope windstorms of san Diego county. Part I: a case study, Mon. Weather Rev. 144 (2016) 529–552.
[17] C. Jones, F. Fujioka, L.M.V. Carvalho, Forecast skill of synoptic conditions associated with Santa Ana winds in Southern California, Mon. Weather Rev. 138 (2010) 4528–4541.
[18] T. Rolinski, S.B. Capps, R.G. Fovell, Y. Cao, B.J. D'Agostino, S. Vanderburg, The Santa Ana wildfire Threat index: methodology and operational implementation, Weather Forecast. 31 (2016) 1881–1897.
[19] M. Billmire, N.H.F. French, T. Loboda, R.C. Owen, M. Tyner, Santa Ana winds and predictors of wildfire progression in southern California, Int. J. Wildland Fire 23 (2014) 1119–1129.
[20] Y. Jin, J.T. Randerson, N. Faivre, S. Capps, A. Hall, M.L. Goulden, Contrasting controls on wildland fires in Southern California during periods with and without Santa Ana winds, J. Geophys. Res. Biogeosciences. (2014).
[21] P.J. Bartlein, S.W. Hostetler, S.L. Shafer, J.O. Holman, A.M. Solomon, Temporal and spatial structure in a daily wildfire-start data set from the western United States (198696), Int. J. Wildland Fire 17 (2008) 8–17.
[22] N.R. Faivre, Y. Jin, M.L. Goulden, J.T. Randerson, Spatial patterns and controls on burned area for two contrasting fire regimes in Southern California, Ecosphere 7 (2016), e01210.

[23] C.A. Kolden, J.T. Abatzoglou, Spatial distribution of wildfires ignited under katabatic versus non-katabatic winds in Mediterranean southern California USA, Fire 1 (2018) 19.

[24] J.G. Edinger, R.A. Helvey, D. Baumhefner, Surface wind patterns in the Los Angeles Basin during "Santa Ana" conditions, Los Angeles, CA, 1964.

[25] M.A. Fosberg, A Case Study of the Santa Ana Winds in the San Gabriel Mountains, Pacific SW Forest and Range Experiment Station, U.S. Forest Service Research Paper PSW-78, Berkeley, CA, 1965.

[26] M.A. Moritz, T.J. Moody, M.A. Krawchuk, M. Hughes, A. Hall, Spatial variation in extreme winds predicts large wildfire locations in chaparral ecosystems, Geophys. Res. Lett. 37 (2010) L04801.

[27] S.H. Peterson, M.A. Moritz, M.E. Morais, P.E. Dennison, J.M. Carlson, Modelling long-term fire regimes of southern California shrublands, Int. J. Wildland Fire 20 (2011) 1–16.

[28] H.K. Preisler, K.L. Riley, C.S. Stonesifer, D.E. Calkin, W.M. Jolly, Near-term probabilistic forecast of significant wildfire events for the Western United States, Int. J. Wildland Fire (2016).

[29] J.H. Scott, M.P. Thompson, J.W. Gilbertson-Day, Exploring how alternative mapping approaches influence fireshed assessment and human community exposure to wildfire, GeoJournal 82 (2017) 201–215.

[30] A. Dye, K. Riley, J. Kim, An evaluation of alternative weather inputs on fire risk simulations in southern California, in: Proc. 6th Int. Fire Behav. Fuels Conf., International Association of Wildland Fire, Albuquerque, NM, 2019.

[31] M.A. Finney, An overview of FlamMap modeling capabilities, in: P. Andrews, B.W. Butler (Eds.), Fuels Manag. To Meas. Success Conf. Proceedings, Portland, OR, March 28-30, 2006, Proceedings RMRS-P-41, USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO, 2006, p. 809.

[32] M.A. Finney, FARSITE: Fire Area Simulator - Model Development and Evaluation, USDA Forest Service Rocky Mountain Research Station, 2004. Research Paper RMRS-RP-4.

[33] M.A. Finney, I.C. Grenfell, C.W. McHugh, R.C. Seli, D. Trethewey, R.D. Stratton, S. Brittain, A method for ensemble wildland fire simulation, Environ. Model. Assess. 16 (2011) 153–167.

[34] M.A. Finney, C.W. McHugh, I.C. Grenfell, K.L. Riley, K.C. Short, A simulation of probabilistic wildfire risk components for the continental United States, Stoch, Environ. Res. Risk Assess. 25 (2011) 973–1000.

[35] J.M. Forthofer, Modeling Wind in Complex Terrain for Use in Fire Spread Prediction, Colorado State University, 2007.

[36] A.L. Sullivan, I.K. Knight, Estimating error in wind speed measurements for experimental fires, Can. J. For. Res. 31 (2001) 401–409.

[37] T. Rolinski, S.B. Capps, W. Zhuang, Santa Ana Winds: A Descriptive Climatology, Weather Forecast, 2018, pp. 257–275.

[38] K.C. Short, M.A. Finney, J.H. Scott, J.W. Gilbertson-Day, I.C. Grenfell, Spatial Dataset of Probabilistic Wildfire Risk Components for the Conterminous United States, CO: Forest Service Research Data Archive, Fort Collins, 2016.

[39] S. Archibald, C.E.R. Lehmann, J.L. Gomez-Dans, R.A. Bradstock, Defining pyromes and global syndromes of fire regimes, Proc. Natl. Acad. Sci. Unit. States Am. 110 (2013) 6442–6447.

[40] A.M. Gill, Fire and the Australian flora: a review, Aust. For. 38 (1975) 4–25.

[41] J. Zachariassen, K. Zeller, N. Nikolov, T. McClelland, A review of the forest Service Remote automated weather station (RAWS) network. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.231.6247&rep=rep1&type=pdf%0Afile:///C:/Users/jcronan.000/Documents/UW/JFSP_ASF/ASF_Info_LibraryZachariassen_et al_2003-0730791169/Zachariassen_etal_2003.pdf%5Cnhttp://190522035340190 3873clipping.bmp, 2003.

[42] K.L. Riley, R.A. Loehman, Mid-21stcentury climate changes increase predicted fire occurrence and fire season length, Northern Rocky Mountains, United States, Ecosphere 7 (2016), e01543.

[43] K.L. Riley, M.P. Thompson, J.H. Scott, J.W. Gilbertson-Day, A model-based framework to evaluate alternative wildfire suppression strategies, Resources 7 (2018).

[44] K.C. Short, A spatial database of wildfires in the United States, 1992-2011, Earth Syst. Sci. Data 6 (2014) 297–366.

[45] R Core Team, R, A language and environment for statistical computing. https://www.r-project.org/, 2017.

[46] B. Everitt, Cluster Analysis, Heinemann Educational Books Ltd., London, 1974.

[47] T.M. Kodinariya, P.R. Makwana, Review on determining number of clusters in K-Means clustering, Int. J. Adv. Res. Comput. Sci. Manag. 1 (2013) 90–95.

[48] WIMS, Chapter 6-Working with Station Information, Boise, ID, 2011.

[49] M.A. Moritz, J.E. Keeley, E.A. Johnson, A.A. Schaffner, Testing a basic assumption of shrubland fire management: how important is fuel age? Source Front. Ecol. Environ. 2 (2004) 67–72. http://www.jstor.org/stable/3868212.

[50] J.E. Keeley, H. Safford, C.J. Fotheringham, J. Franklin, M. Moritz, The 2007 southern California wildfires: lessons in complexity, J. For. 107 (2009) 287–296.

[51] P.A.M. Fernandes, Fire spread prediction in shrub fuels in Portugal, For. Ecol. Manage. 144 (2001) 67–74.

[52] R.E. Clark, A.S. Hope, S. Tarantola, D. Gatelli, P.E. Dennison, M.A. Moritz, Sensitivity analysis of a fire spread model in a chaparral landscape, Fire Ecol 4 (2008) 1–13.

[53] M.A. Moritz, Spatiotemporal analysis of controls on shrubland fire regimes: age dependency and fire hazard, Ecology 84 (2003) 351–361.

[54] J.T. Abatzoglou, J.K. Balch, B.A. Bradley, C.A. Kolden, Human-related ignitions concurrent with high winds promote large wildfires across the USA, Int. J. Wildland Fire 27 (2018) 377–386.