CSIRO PUBLISHING

RESEARCH PAPER
https://doi.org/10.1071/WF21177

International Journal of WILDLAND FIRE
International Association of Wildland Fire

Human- and lightning-caused wildland fire ignition clusters in British Columbia, Canada

Sean C. P. Coogan[A,*], Olivia Aftergood[A] and Mike D. Flannigan[B]

For full list of author affiliations and declarations see end of paper

*Correspondence to:
Sean C. P. Coogan
Department of Renewable Resources, University of Alberta, Edmonton, AB T6G 2H1, Canada
Email: scoogan@ualberta.ca

Received: 7 December 2021
Accepted: 12 September 2022
Published: 11 October 2022

Cite this:
Coogan SCP *et al.* (2022)
International Journal of Wildland Fire
31(11), 1043–1055. doi:10.1071/WF21177

ABSTRACT

Wildland fire is a common occurrence in western Canada, with record-setting area burned recorded in British Columbia (BC) in the past decade. Here, we used the unsupervised machine learning algorithm HDBSCAN to identify high-density clusters of both human- and lightning-caused wildfire ignitions in BC using data from 2006 to 2020. We found that human-caused ignition clusters tended to occur around population centres, First Nations communities, roads and valleys, and were more common in the southern half of the province, which is more populated. Lightning-ignition clusters were generally fewer in number and larger in size than human-caused fires for most hyperparameter settings. There were significant differences ($X^2 = 1884.8$, d.f. = 7, P-value $< 2.2 \times 10^{-16}$) in fuels associated with lightning- versus human-caused ignition clusters, with human-ignition cluster fires being more often found within leafless aspen (D1) and ponderosas pine and Douglas fir (C7) fuel types. These high-density clusters highlight regions where the greatest densities of both lightning- and human-caused fires have occurred in the province, thereby identifying regions of potential interest to wildland fire managers, researchers and various communities and industries.

Keywords: Canada, clustering, fuels, HDBSCAN, human-caused fires, interface fires, lightning-caused fires, unsupervised machine learning.

Introduction

Wildland fire has been a longstanding feature of Canadian forests, where many species of flora and fauna have evolved with fire as a recurring disturbance (Coogan *et al.* 2021). Although wildland fire provides many positive benefits to Canadian forests, it can obviously have significant (even catastrophic) impacts when it affects communities and infrastructure (Coogan *et al.* 2019). Wildfires in Canada have burned an average of 1.96 Mha per annum based on data from 1959 to 2015, and there has been a significant increase in the number of large wildfires (≥200 ha) and area burned during this time, possibly due to increases in lightning-caused fires (Hanes *et al.* 2019). The main ignition sources of wildfires in Canada are lightning and people, each accounting for ~50% of ignitions nationwide (Stocks *et al.* 2002; Hanes *et al.* 2019; Coogan *et al.* 2020). The number of human-caused fires in Canada, which account for approximately 10% of total area burned, appears to have decreased since 1959 (Hanes *et al.* 2019). Human-caused fires, however, often occur during the shoulder seasons, which has resulted in an increase in the length of the fire season in several regions in Canada and elsewhere (Balch *et al.* 2017; Hanes *et al.* 2019).

Human-caused wildland fires occur most often in interface areas that are in close proximity to where people reside or have access (Pew and Larsen 2001; Gralewicz *et al.* 2012; Price and Bradstock 2014; Johnston and Flannigan 2018). The wildland–urban interface (WUI) is a term used to describe areas where people and their developments are in contact with, or are dispersed within, wildland fuels (USDA and USDI 2001). Despite the name, the WUI also includes areas that may not be considered urban per se such as small population centres, reservations and holiday communities (Johnston and Flannigan 2018).

Expanded definitions of the WUI have been developed in Canada to include the wildland–industrial and wildland–infrastructure interfaces, which all together can be considered the wildland–human interface (Robinne *et al.* 2016; Johnston and Flannigan 2018). Lightning-caused fires tend to occur in wildland–human interface areas in lower proportions than human-caused fires and often occur in more remote areas of Canada where human access is limited (Johnston and Flannigan 2018). Lightning-caused fires in remote areas are often left to burn freely for positive ecological benefits (Tymstra *et al.* 2020). On the other hand, lightning-caused fires can also be associated with high suppression costs and have a greater chance of becoming large fires than human-caused fires, which tend to occur closer to population centres where detection may be quicker and fire suppression more aggressive (Stocks *et al.* 2002).

Given that human-caused fires tend to occur in areas close to human population centres, developments and recreational areas, it is likely that such fires are geographically clustered. Likewise, lightning-caused fires may also be geographically clustered, likely in relation to the distribution, type and receptivity (i.e. moisture content) of wildland fuels and topography, as well as patterns of lightning-weather activity (Nadeem *et al.* 2020). However, such clusters may not be readily apparent. Identifying ignition clusters is important for characterising the spatial distribution of wildland fire, which is of direct importance to wildland fire management and suppression planning, including for the prevention and preparedness phases. Furthermore, a better understanding of ignition clusters may lead to ecological insights given the prominent role fire can play in ecosystems (Coogan *et al.* 2021).

Recently, researchers have suggested that machine learning may be useful for further understanding the intricacies of wildland fire (Jain *et al.* 2020). In this paper, we explore an unsupervised machine learning algorithm to identify geographic high-density wildfire ignition clusters for both human- and lightning-caused fires in British Columbia (BC), Canada. The province of BC has experienced noteworthy fire activity in recent years, with record-setting area burned in 2017 and 2018 (Government of BC 2021). The province also contains among the highest amount of WUI areas in Canada (Johnston and Flannigan 2018). More recently, the 2021 heat wave prompted another active fire season in BC in which the town of Lytton was largely destroyed by wildfire (Schiermeier 2021). Extreme fire weather has increased over time in BC (Jain *et al.* 2021), and more challenging fire seasons are expected in the future owing to climate change (Wang *et al.* 2017; Wotton *et al.* 2017).

In this paper, we focus on clustering wildfires in BC using the density-based algorithm HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise; Campello *et al.* 2013; McInnes *et al.* 2017) to identify areas with the highest densities of both lightning- and human-caused fires. We also compared the fuel types associated with clusters in the province. The regions associated with the clusters identified using HDBSCAN are likely to be of interest to wildland fire managers, researchers, ecologists, communities and industries, among other interested parties in the province and beyond. This research is also likely to be of interest to those interested in unsupervised machine learning clustering as well as those wishing to understand the spatial clustering of human- and lightning-caused wildland fire events more generally.

Materials and methods

Study area

BC contains a variety of ecosystems including temperate rainforest, semideserts, grasslands, boreal forest and alpine tundra, depending on such factors as proximity to the Pacific Ocean, latitude and topography (Meyn *et al.* 2010). Mountains comprising part of the North American Cordillera run from south to north through BC and restrict westward airflow, resulting in rain-shadow effects east of the mountains ranges including in parts of Alberta (Wierzchowski *et al.* 2002; Moore *et al.* 2010).

The different characteristics of the various ecosystems in the province, including canopy cover and surface fuel properties, influence various wildland fire ignition factors such as fuel moisture wetting and drying rates and thus fuel receptivity to ignition. The moisture content and receptivity of fuels in western Canada is strongly influenced by the mid-latitude storm track, which brings high-moisture air from the Pacific to western Canada. In the summer (approximately June to August), the North Pacific High often blocks the flow of this high-moisture air, resulting in warm and dry conditions that can last many weeks and increase the dryness of both surface and deep organic fuel layers (Nadeem *et al.* 2020). In fact, climate change may cause an increase in blocking ridges that persist for long periods in western Canada, thereby creating more dangerous wildland fire conditions (e.g. Petoukhov *et al.* 2018).

Lightning occurrence is highly episodic and mostly occurs during dominant storm events. Across Canada, cloud-to-ground lightning occurs primarily from June to August (Burrows and Kochtubajda 2010). Similarly, the majority of lightning-caused fires ≥ 2 ha in BC occur in August (Coogan *et al.* 2020). Regarding lightning-ignition efficiency, 1 in 50 (2%) lightning discharges have been shown to start a fire in BC's central Cordillera (Wierzchowski *et al.* 2002).

Population centres are primarily located in the southern half of the province (Supplementary Material S1, Supplementary Fig. S1). Outdoor recreation and camping are popular in BC, with some regions experiencing high numbers of tourists. Human-caused fires tend to occur primarily in the spring (during April and May; Coogan *et al.* 2020). Because the bulk of human-caused fires tends to occur during the shoulder seasons (i.e. primarily in the spring, but also in the autumn) outside the typical lightning-caused fire season, they extend the length of the fire season overall (Hanes *et al.* 2019; Coogan *et al.* 2020).

Data

Fire data

We used point data from the Canadian National Fire Database (CNFDB; Canadian Forest Service 2021) for our analysis. The CNFDB point data are a dataset of wildfire locations provided collaboratively by Canadian fire management agencies and are available online. The database has some limitations, however, as it may include errors, and data completeness and quality may vary over time and among contributing management agencies – see Hanes *et al.* (2019) for a discussion on the strengths and limitations of Canadian fire databases. At the time of our analysis, the CNFDB point data were available up to and including the year 2020. We did not include Parks Canada agency data in our analysis, because many of those fires did not include specific geographic coordinates. We focused our analysis on a 15-year period from 2006 to 2020: we reasoned that such a period would include data recorded and collected using similar and consistent methods, as procedures have changed over longer time periods. Furthermore, there has been an increase in extreme fire weather globally during this period (Jain *et al.* 2021). We clustered fires of all sizes because we were interested in identifying clusters of fire ignitions regardless of their final size. The subsets of human- and lightning-caused raw data ignition points we used in our analysis are displayed in Supplementary Material S1, Supplementary Figs S2, S3. The wildfire ignition points were clustered based on their geographic location, where the latitudes and longitudes of the spatial points were transformed to radians for input into the haversine distance calculation before clustering.

Other data

We used a basemap and anthropogenic data to help put the geographic clustering of human- and lightning-caused fires in context with landscape features and characteristics. For instance, we used road data from the CanVec series published by Natural Resources Canada (https://open.canada.ca/data/en/dataset/8ba2aa2a-7bb9-4448-b4d7-f164409fe056) and population centre data from Statistics Canada (https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2016-eng.cfm) (Supplementary Material S1, Supplementary Fig. S1). Figures and clustering were created using Python version 3.85, with the exception of fuels raster maps and chi-squared tests, which were produced using R version 4.1.0 (see below). We used the *matplotlib* plotting library (Caswell *et al.* 2021) to create figures and the *contextily* package (https://contextily.readthedocs.io/en/latest/) for the basemaps; note that the *contextily* basemaps used the Web Mercator projection (EPSG:3857). To evaluate the relationships between clustering and fuel types, we used the aggregated (40-km pixel size) Forest Fire Behaviour Prediction (FBP) System fuel type raster classification (Supplementary Material S1, Supplementary Fig. S4) published in Wotton *et al.* (2017). This larger-scale raster is based on the assumption that fire will burn the more volatile fuels in a raster cell. Significant differences in fuel types associated with cluster and non-cluster fire points were evaluated using the chi-squared test.

HDBSCAN clustering

About HDBSCAN

There were several reasons we used the HDBSCAN algorithm to identify high-density clusters of wildfire ignition points in BC. HDBSCAN assigns high-density regions to clusters while leaving out noise (i.e. 'outliers') that does not belong to a particular cluster – compare this with other partitioning clustering algorithms, such as *k*-means, which partitions all points in a dataset to a particular cluster. This was useful for our application, as the plots of the raw data (Supplementary Material S1, Supplementary Figs S2, S3) showed that there was extensive coverage of wildfire ignition points across the province that were likely not representative of areas in which a high density of wildfire ignitions had occurred. HDBSCAN also works well for arbitrarily shaped clusters (as opposed to algorithms that assume spherical clusters), which is useful for identifying clusters along valleys and roads. HDBSCAN also offers an improvement over the original DBSCAN algorithm (Ester *et al.* 1996) by identifying clusters of different sizes and densities. HDBSCAN also allows for the use of the haversine distance metric (Robusto 1957), important in navigation, which determines the great-circle distance between two points on a sphere given their latitudes and longitudes (the haversine metric thus assumes that the Earth is a sphere). The haversine metric may provide a more realistic distance metric for clustering geolocation data across larger distances than Euclidean distance (Maria *et al.* 2020). HDBSCAN also automatically selects the optimal number of clusters for a particular dataset, which is advantageous for our unlabelled dataset in which the 'true' clusters are not known – this is also an advantage over some other popular algorithms that require the number of clusters to be given *a priori*.

In short, the basic mechanics of HDBSCAN follows a series of five steps to develop clusters (https://hdbscan.readthedocs.io/en/latest/how_hdbscan_works.html). In the first step, HDBSCAN transforms the space according to the density of points. Second, a minimum spanning tree of a distance-weighted graph is constructed, followed by the third step, in which HDBSCAN builds a cluster hierarchy of connected components. In the fourth step, HDBSCAN condenses the cluster hierarchy based on the minimum cluster size (hereafter 'min_cluster_size') hyperparameter (which is set by the user), and, in the fifth step, extracts stable clusters from the condensed tree.

Hyperparameter tuning and selection

There are two main hyperparameters to set when using HDBSCAN, namely min_cluster_size and min_samples. Additional advanced parameters can be set in HDBSCAN; however, we do not delve into these options in this paper. For more information on these hyperparameters, we refer the reader to the HDBSCAN Clustering Library (https://hdbscan.readthedocs.io/en/latest/index.html).

The primary parameter affecting clustering results is min_cluster_size, which is the minimum number of points that a final cluster can contain. There is no objective method for selecting min_cluster_size: it is chosen based on domain-specific knowledge and experimentation. The HDBSCAN documentation states that ideally min_cluster_size is a fairly intuitive parameter to select; however, this may not always be the case in practice. In general, the higher min_cluster_size is, the larger clusters will be. Likewise, setting min_cluster_size lower will generally result in a greater number of smaller clusters.

The hyperparameter min_samples is the minimum number of neighbours to a core point. In essence, min_samples provides a way to adjust how conservative the model will be. The larger the value of min_samples (which cannot exceed min_cluster_size), the more conservative clustering will be – more points will be assigned as outliers (noise) and the clusters will be restricted to increasingly dense areas. The default min_samples setting is the same as min_cluster_size because this is considered to be a reasonable default; however, the min_samples parameter can be adjusted if users are having problems with clustering.

Advanced adjustments to both min_cluster_size and min_samples can be made depending on the desired clustering outcome. For example, three basic clustering strategies may include: (1) using a small min_samples and a small min_cluster_size for many highly specific clusters; (2) using a small min_samples and a large min_cluster_size for more generalised clusters with high detail; (3) using a large min_samples and a large min_cluster size for very general clusters with a lot of noise (outliers) removed.

In our paper, we focused on optimising the hyperparameter min_cluster_size where we set the search space to contain a range of min_cluster_size values in increments of 25, from 25 to 200 (i.e. 25, 50, 75, 100, 125, 150, 175 and 200) and left min_samples at the default value (i.e. the same as min_cluster_size). After evaluating the results of varying min_cluster_size, we explored varying min_samples, the outcomes of which are described in the *Results* section.

A note should be made here about our philosophy regarding the final hyperparameter selection for clustering. As described above, we chose to select hyperparameter settings and clustering results based on our expert opinion, where our primary goal was to choose the hyperparameter settings that balanced the trade-off between generating many very small localised clusters and fewer larger, more generalised clusters at the provincial scale. Clustering itself can be a subjective exercise where there is no definition of what a 'true' cluster is (Hennig 2015), especially for unlabelled datasets such as ours. Methods exist for evaluating various clustering outcomes; however, many of these common indices are not appropriate for density-based clustering. The Density-Based Clustering Validation (DBCV) method can be used to evaluate HDBSCAN clustering outcomes, but this method also has limitations (Moulavi *et al.* 2014). Experts have been considered to be more consistent in evaluating clusters than cluster validation indices (Lewis 2009), and this is the approach we took for this paper. Given the subjective nature of our approach, the full clustering results of our hyperparameter exploration are presented in the Supplementary Material for transparency and to allow the reader to evaluate the sensitivity of the algorithm to the hyperparameter tuning. Given the results we obtained, we do not consider this to be an inappropriate or controversial approach to our analysis.

Results

Human-caused fires clusters

We found that the number of human-caused fire clusters decreased as we increased min_cluster_size from 25 to 200, ranging from 58 to 6 respectively (Fig. 1). Although there was a fairly large range in the number of clusters produced, the clusters themselves were generally found in similar locations, whether smaller and more localised or larger and more generalised. To illustrate, we display the results of clustering with min_cluster_sizes of 25 and 200 in Fig. 1. The full range of min_cluster_size figures can be found in the Supplementary Material S1, Supplementary Figs S5–S12.

We explored varying min_samples for human-caused fires; however, we did not find that this exercise added any more useful information to what we had already achieved by varying min_cluster_size (i.e. clusters occurred in the same general areas but varied in number and size). Thus, we kept min_samples at the default setting and considered only different min_cluster_size settings.

For our analysis of human-caused fires at the provincial scale, we chose to use a min_cluster_size of 75. One reason we chose this hyperparameter setting is that in larger settings, the clusters on Vancouver Island joined together with the cluster on the mainland around the Whistler area. Likewise, at settings >125, we found that the clusters started to become too large in the sense that they covered relatively large areas of the province that contained several different communities (similar to the clusters formed using a min_cluster_size of 200 in Fig. 1). We did, however, consider that smaller min_cluster_size values could produce useful results at smaller scales. To illustrate this, we examined

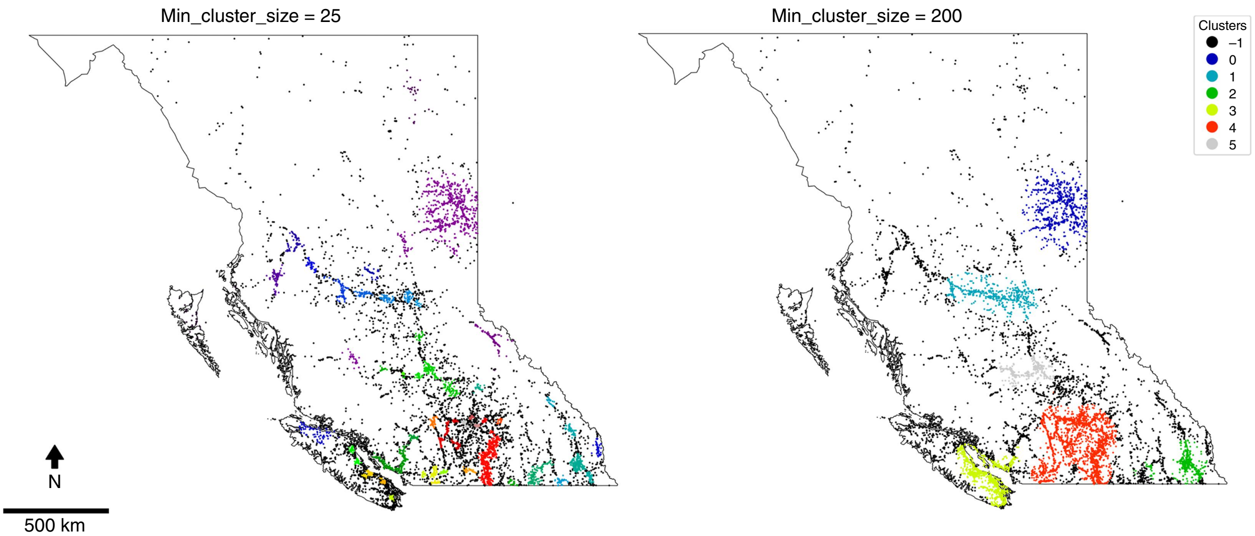

Fig. 1. HDBSCAN clustering results of human-caused fires (2006–2020) in British Columbia, Canada, for min_cluster_sizes of 25 and 200. Note that HDBSCAN assigns points that do not belong to clusters as −1 (black points) and cluster identification numbers start at 0. The legend for min_cluster_size 25 is not shown because there were a total of 58 clusters identified and these would not fit in the image. Furthermore, because of the large number of clusters displayed in the min_cluster_size = 25 figure, the variation in colours between clusters is very subtle in many cases.

the clusters generated using min_cluster_size of 25 centred around the Kamloops area, which is discussed below.

Using a min_cluster_size of 75, we found a total of 17 clusters among the 9503 human-caused ignition points in the province of BC from 2006 to 2020 (Fig. 2). Note that in the Python implementation of HDBSCAN, cluster numbering starts at 0 instead of 1, and points labelled '−1' have been identified as 'outliers' and are not part of a cluster; we maintained this original numbering system. As predicted, most human-ignition clusters in BC were associated with roads and population centres (Table 1), as were many non-cluster outlier points. Some clusters were associated with population centres with <1000 inhabitants (e.g. clusters 0 and 8), which were not symbolised on our maps. Large population centres with >100 000 residents were also found within or adjacent to clusters (e.g. Kelowna in cluster 11 and Abbotsford in cluster 9). First Nations communities were likewise found within some human-caused fire clusters such as clusters 3, 8 and 14. Furthermore, Provincial Parks and popular recreation areas were associated with clusters (e.g. cluster 0). Topography also seemed to play a critical role in the location of human-caused fires clusters, with many clusters being located in and around valleys (e.g. clusters 0, 5 and 11).

As mentioned previously, we considered that using a min_cluster_size of 25 would be useful for examing clustering at a more localised scale. Thus, we examined clustering around the Kamloops region of BC, which has experienced noteworthy fire activity across the time scale of our data. Twelve distinct clusters formed around population centres of all sizes in the region, such as Lytton (<1000 residents), Merritt, Kamloops and Kelowna (Fig. 3). Clusters also formed around First Nations Reservations such as the Skeetchestn and Nooaitch Indian Bands. The largest cluster formed around the Okanagan Valley region, which included several population centres. Roads and topography (valleys) seemed to play a prominent role in the formation of the human-caused wildfire clusters. The majority of clusters using a min_cluster_size of 25 seemed to be part of larger clusters identified using a min_cluster_size of 75; however, some new clusters emerged at the lower setting such as the Chase, Lillooet and Sicamous clusters.

Lightning-caused fire clusters

When we varied the min_cluster_size of lightning-caused fires (n = 14 759), we found that fewer clusters were produced than for human-caused fires, with most forming three clusters; however, two clusters were formed when min_cluster size was set to 50 and 200 (Supplementary Material S1, Supplementary Figs S13–S21). The pattern of clustering stabilised at a min_cluster_size of 75, where the clustering did not change much until 200, where the Vancouver Island cluster joined the mainland cluster. Interestingly, the dominant three-cluster pattern seems to have a relationship with ecozones in the province, with two of the three clusters showing a close overlap with the Taiga Plains and Montane Cordillera, as demonstrated in Fig. 4 in which we plot the clustering results for min_cluster_size 75.

We varied min_samples for lightning-caused fires using a grid search approach, in which we varied min_samples by 1, 5, 10, 15 20, and 25, as well as the default (i.e. the same

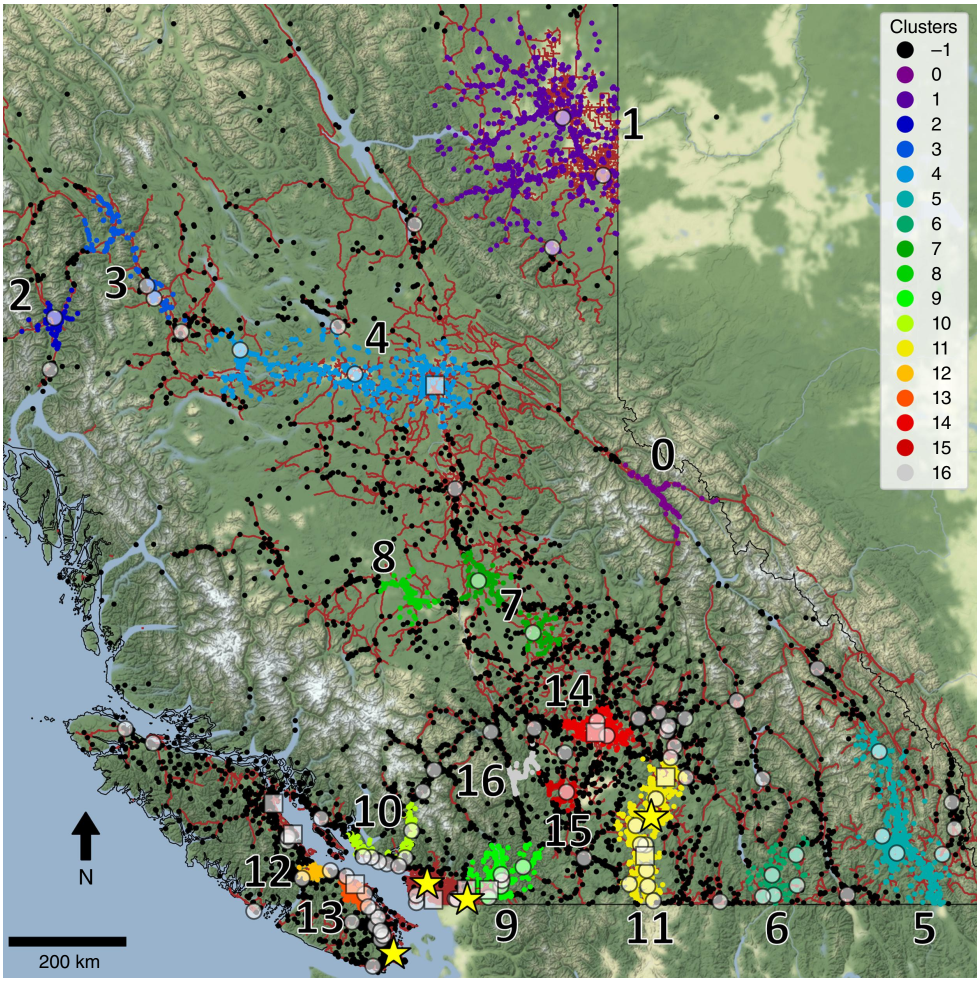

Fig. 2. HDBSCAN clusters of human-caused ignition points (2006–2020) in British Columbia, Canada. A total of 17 clusters were identified (numbered 0–16). Ignition points labelled −1 (black) were identified as 'outliers' and are not considered part of a cluster. Population centres are shown as circles (1000–29 999), squares (30 000–99 999) and stars (≥100 000). The names of population centres associated with clusters are given in Table 1. Points are plotted on top of a base map with roads (brown lines). Basemap tiles provided by Stamen Design, CC BY 3.0 – Map data (C) OpenStreetMap contributors. Map projection: Web Mercator (EPSG:3857).

value as min_cluster_size) for each min_cluster_size setting (Supplementary Material S1, Supplementary Table S1, Supplementary Material S2). For min_cluster_size 25, we found a large number of clusters produced for min_samples of 1, 5 and 10 (189, 145, and 107 clusters, respectively); however, clusters returned to three for higher min_samples settings. The mode number of clusters produced across our grid search was three (n = 26). The second most common number of clusters produced was 2 (n = 8) followed by 5 and 10 (n = 4 each). We found that settings that produced five clusters tended to form generalisable groups of the more fragmented 10-cluster maps. Furthermore, a min_samples setting of 10 for both min_cluster_size 125 and 150 produced very similar maps with five clusters. We ultimately chose a min_cluster_size of 125 with a min_samples of 10 for our analysis of associations with fuels (Fig. 5). This clustering resulted in similar clusters in the Taiga Plains and Vancouver Island as found in most of our clustering maps; however, the larger Montane Cordillera cluster was split into three separate clusters, with two of the smaller clusters forming in central BC.

Similarly to our human-caused fire clustering, we also explored lightning-caused fire clustering at a low density hyperparameter setting (i.e. min_cluster_size = 25, min_samples 10) to examine clustering at a more localised scale. This setting produced a large number of clusters province-wide (n = 107); however, we focused on a relatively small subset of clusters around the Kamloops area (Fig. 6) to draw comparisons with our human-caused fire clustering. In this region, more lightning-caused fire clusters were found away from population centres than human-caused fire clusters; however, some lightning-fire cluster

Table 1. General population centres and regions associated with human-caused wildfire clusters in British Columbia, Canada, from 2006 to 2020.

Cluster number	Population centres or region
0	Valemount,[A] McBride,[A] Mt Robson Provincial Park
1	Fort St John, Dawson Creek, Tumbler Ridge
2	Terrace
3	Smithers, Telkwa, Hazelton,[A] Gitanyow[B]
4	Prince George, Vanderhoof, Fort St James, Burns Lake
5	Kimberley, Cranbrook, Fernie, Invermere
6	Nelson, Salmo, Rossland, Trail, Castlegar
7	Williams Lake, 100 Mile House
8	Alexis Creek,[B] Redstone[B]
9	Abbotsford, Chilliwack, Hope
10	Sechelt, Gibsons, Squamish, Whistler
11	Okanagan Valley area (e.g. Vernon, Kelowna, Penticton, Osoyoos)
12	Parksville, Port Alberni
13	Nanaimo
14	Kamloops
15	Merritt
16	Lytton

[A]<1000 residents and therefore not shown on the map.
[B]First Nations community or unincorporated settlement.

areas overlapped with human-cluster areas such as around Kamloops. Some lightning-fire clusters were also found along valleys and roads, while others were in proximity to water bodies.

Fuels associated with clusters

We found that the FBP System fuel types associated with human ignition cluster points were statistically different from human ignition outlier (non-cluster) points (X^2 = 1058.4, d.f. = 8, P-value $<2.2 \times 10^{-16}$; Table 2). The chi-square results were most influenced by differences between the ponderosas pine and Douglas fir (C7) and leafless aspen (D1) fuel types, where a higher proportion of the cluster points were associated with these fuels than the outlier points; this is demonstrated graphically in a correlation plot of the standardised residuals (Supplementary Material S1, Supplementary Fig. S22). A small number of points were found in the 'no fuels' category, likely due to the larger raster aggregation and to error associated with the fire point locations. However, the number of fires in the non-fuel category was relatively small and seemingly acceptable for our purposes (0.1–2.3% for all results in Table 2); therefore, the results should still provide a useful generalisation. FBP System fuel types associated with lightning-caused ignition cluster points were statistically different from non-cluster lightning ignition points (X^2 = 1688.6, d.f. = 8, P-value $<2.2 \times 10^{-16}$; Table 2), where the chi-square results were strongly influenced by differences in the mature jack or lodgepole pine (C3) fuel type (Supplementary Material S1, Supplementary Fig. S23). In fact, no cluster points of either

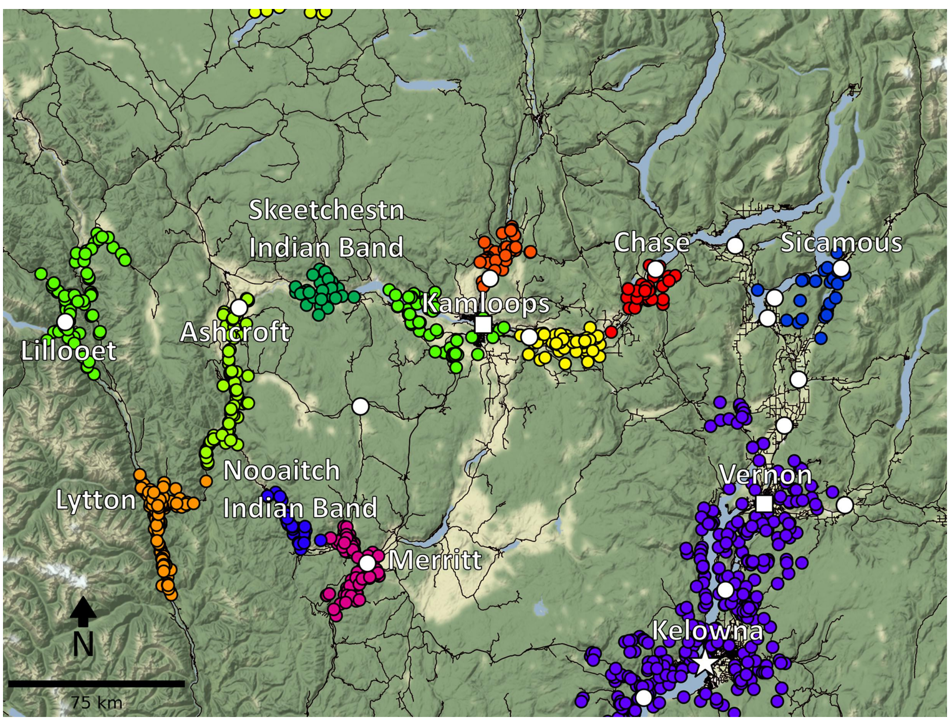

Fig. 3. Human-caused fire clusters generated using a min_cluster_size setting of 25 for the Kamloops region in BC. Points in a cluster have the same colour. Non-cluster points (i.e. outliers) have been removed for clarity. Roads are shown as black lines. Population centres are shown as circles (1000–29 999), squares (30 000–99 999) and stars (≥100 000). Lytton had <1000 residents and has no symbol on the map. Likewise, the First Nations Reservations were not included in the population centre data. Basemap tiles provided by Stamen Design, CC BY 3.0 – Map data (C) OpenStreetMap contributors. Map projection: Web Mercator (EPSG:3857).

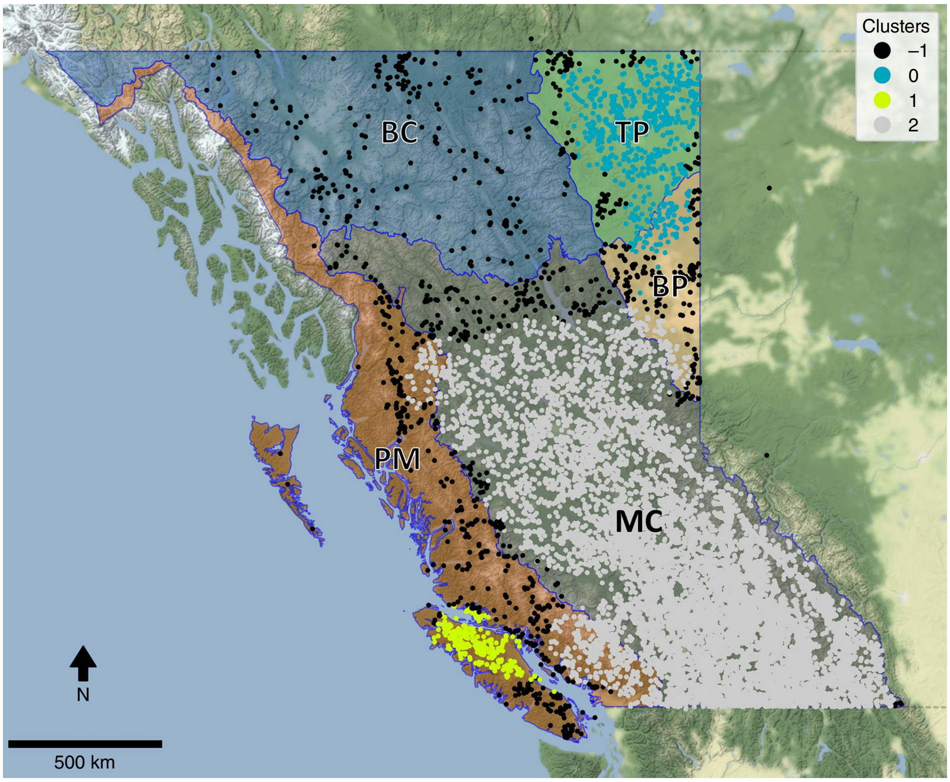

Fig. 4. HDBSCAN clusters (min_cluster_size = 75) of lightning-caused ignition points (2006–2020) in British Columbia, Canada, plotted on top of Canadian ecozones including the Montane Cordillera (MC), Boreal Plains (BP), Taiga Plains (TP), Boreal Cordillera (BC) and Pacific Maritime (PM). Basemap tiles by Stamen Design, CC BY 3.0 – Map data (C) OpenStreetMap contributors. Map projection: Web Mercator (EPSG:3857).

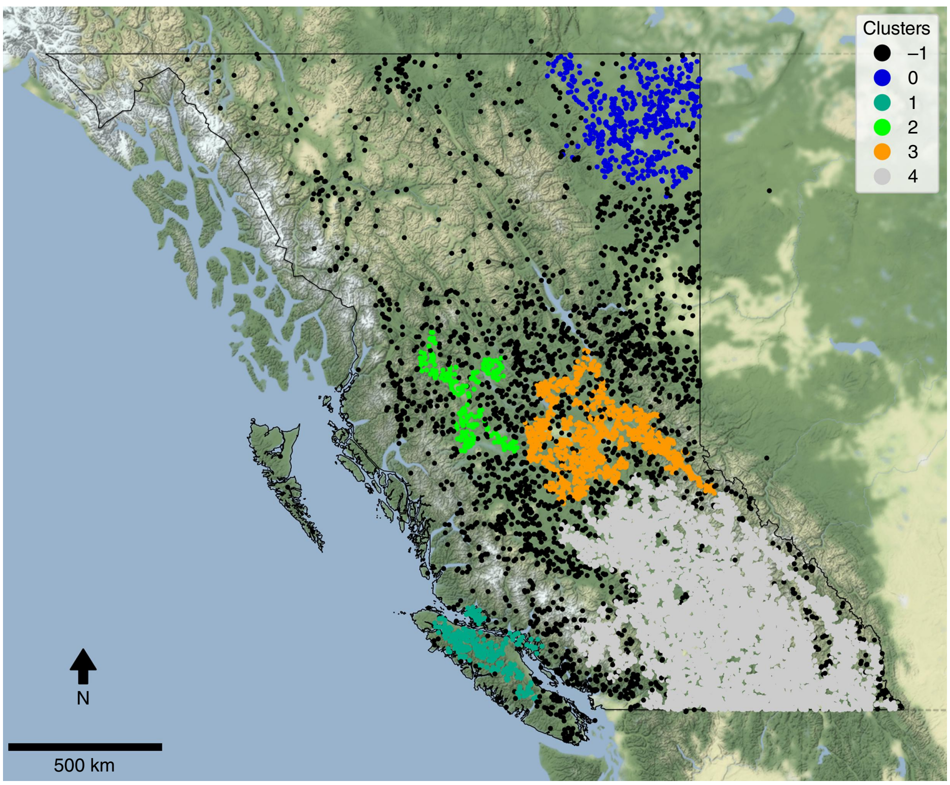

Fig. 5. HDBSCAN clusters (min_cluster_size = 125, min_samples = 10) of lightning-caused ignition points (2006–2020). Basemap tiles by Stamen Design, CC BY 3.0 – Map data (C) OpenStreetMap contributors. Map projection: Web Mercator (EPSG:3857).

lightning- or human-caused fires were found in the C3 fuel type. Lightning and human clusters were significantly different ($X^2 = 1884.8$, d.f. = 7, P-value $< 2.2 \times 10^{-16}$), and the results were driven mostly by the differences in the C7 and D1 fuel types where human-caused fire clusters were found in higher proportions (Supplementary Material S1,

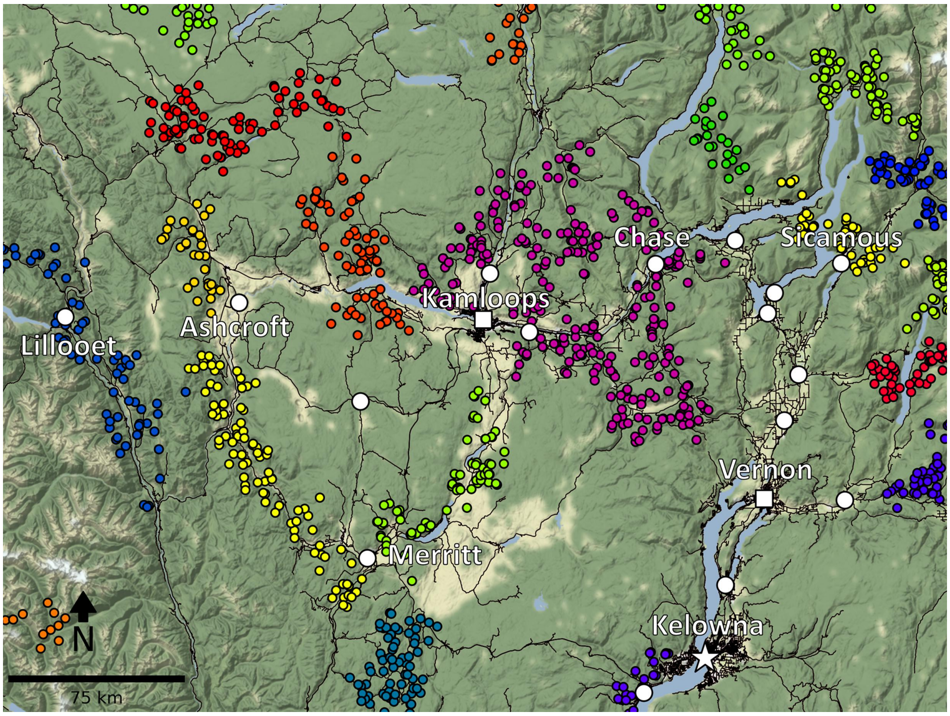

Fig. 6. Lightning-caused fire clusters generated using a min_cluster_size setting of 25 and min_samples setting of 10 for the Kamloops region in BC. Points in a cluster have the same colour. Non-cluster points (i.e. outliers) have been removed for clarity. Roads are shown as black lines. Population centres are shown as circles (1000–29 999), squares (30 000–99 999) and stars (≥100 000). Basemap tiles provided by Stamen Design, CC BY 3.0 – Map data (C) OpenStreetMap contributors. Map projection: Web Mercator (EPSG:3857).

Table 2. The percentage of fuel types associated with wildfires in clusters compared with non-cluster ignitions British Columbia, Canada, for both human-caused and lightning-caused fires.

FBP fuel type	Fuel description	Human-caused fires		Lightning-caused fires	
		Non-cluster (%)	Clusters (%)	Non-cluster (%)	Clusters (%)
Non-fuel	Non-fuel	2.3	0.4	1.1	0.1
C1	Open spruce–lichen woodland	1.1	<0.1	2.2	1.7
C2	Boreal black spruce	25.1	19.5	44.6	29.6
C3	Mature jack or lodgepole pine	1.5	0	7.3	0
C5	Red and white pine	60.7	46.1	39.8	58.1
C7	Open Douglas fir or ponderosa pine	9.0	22.1	0.2	9.1
D1	Leafless aspen	<0.1	7.6	1.4	0
M1	Boreal deciduous and conifer mixedwood with 50% conifer or more	0.4	3.9	3.2	1.3
M1(25)	Boreal deciduous and conifer mixedwood with 25% conifer or less	<0.1	0.3	0.2	0

There were significant differences between human-caused fires in clusters versus non-clusters ($X^2 = 1058.4$, d.f. = 8, *P*-value < 2.2×10^{-16}), lightning-caused fires in clusters versus non-clusters ($X^2 = 1688.6$, d.f. = 8, *P*-value < 2.2×10^{-16}), and between human-caused fire clusters and lightning-caused fire clusters ($X^2 = 1688.6$, d.f. = 8, *P*-value < 2.2×10^{-16}).

Supplementary Fig. S24). C3 fuels were removed from the chi-square analysis of human- and lightning-caused clusters because there were no cluster points in this fuel type for either category. There were also significant differences between fuels associated with all (both cluster and non-cluster points) lightning ignitions versus all human-caused ignitions ($X^2 = 1269.7$, d.f. = 8, *P*-value $< 2.2 \times 10^{-16}$), with a higher proportion of human-caused ignitions in D1 and C7 fuels, and fewer in boreal black spruce (C2) and open spruce–lichen woodland (C1) fuels (Supplementary Material S1, Supplementary Fig. S25). Cluster points are shown plotted on the fuel type raster in Supplementary Material S1, Supplementary Figs S26, S27.

Discussion

An important insight that can be drawn from this paper is that high-density clusters of human-caused fires tend to be associated with population centres and roads. Furthermore, topography seems to have played a role, with many clusters forming in more inhabited valleys. This is consistent with research that has shown that human-caused ignitions in Canada tend to be more frequent in interface areas in which humans live and have access (Wotton *et al.* 2003; Gralewicz *et al.* 2012; Robinne *et al.* 2016; Johnston and Flannigan 2018). Human-caused fires clusters also occurred in areas experiencing high levels of industrial activity, such as the Fort St John cluster, which may have played a role in fire occurrence. Furthermore, human-caused fire clusters were found around some population centres that have fairly high amounts of tourism and outdoor recreation activities such as the Okanagan Valley, Prince George and Mount Robson Provincial Park. Clusters also formed in areas with First Nations communities, which can be at a high risk from wildfire (McGee 2021).

Identifying human-caused ignition clusters such as the ones we did may be a useful initial step in identifying problematic areas and may ultimately lead to a better understanding of the factors driving these human-ignition hotspots. Interestingly, cluster 16 formed around the town of Lytton, which was largely destroyed by a wildfire that started on 30 June 2021, following a day in which the town set the record for the hottest day ever recorded in Canada (49.6°C; Schiermeier 2021). The cause of the Lytton wildfire is at this point undetermined, but it may be due to human causes. Fortunately, human-caused fires appear to be declining over time in BC (Coogan *et al.* 2020). Yet, given that extreme fire weather has increased over time and that climate change is expected to lead to more dangerous fire seasons, and given the increasing development in the wildland–human interface, instances such as this may become more common. Thus, further reducing the amount of preventable human-caused wildfires is important given that climate change may lead to an increased occurrence of lightning-caused fires and further increase the negative impacts of human-caused wildfires on ecosystems and communities, including smoke, which can be transported across large distances.

The prevalence of human-caused wildfires near population centres and in interface areas is not just a Canadian phenomenon, but has been observed all over the world. However, the relationship between human population density and wildland fire is complex, and has been shown to be non-linear in many regions across the world because population centres can offer both sources of ignition and enhanced protection from wildland fire spread owing to increased suppression activity (Bistinas *et al.* 2013; Price and Bradstock 2014). In such cases, the incidence of wildland fire ignitions increases with population density up to a local threshold, then decreases. In our paper, we found that population centres of all sizes were associated with individual clusters, which further supports the notion that the relationships between human-caused wildfires and population density are non-linear and involve other factors.

In our study, the larger lightning-ignition clusters were found more or less in the areas where the highest lightning activity has been identified for the province by Burrows and Kochtubajda (2010), with the exception of the Vancouver Island cluster. The main lightning-fire clusters in BC were located in the central and southern half of the province, associated with the Montane Cordillera ecozone, which is among the most diverse ecozones in the country. The lightning-fire cluster in northeast BC approximately overlaps the Taiga Plains ecozone, which extends down from the Northwest Territories. These clusters were separated by the Boreal Plains ecozone, possibly owing to there being more deciduous fuel in this area: coniferous fuels are associated with more frequent lightning ignitions in this ecozone compared with deciduous fuels like aspen (*Populus* spp.; Krawchuk *et al.* 2006). In fact, the fuels raster we used showed a greater abundance of aspen fuels in the Boreal Plains ecozone compared with the Taiga Plains and Montane Cordillera (Supplementary Material S1, Supplementary Fig. S4), and the points associated with the Taiga Plains cluster tended to be found more in conifer compared with mixedwood fuels (Supplementary Material S1, Supplementary Fig. S27).

Interestingly, we also identified a lightning-ignition cluster on the north of Vancouver Island in BC. Recent research completed on Vancouver Island concerning climate change adaptations and planning has highlighted drought as a main concern for locals. Local managers and residents of the island have been experiencing hotter and drier summers, which have led to drought conditions and increased wildfire activity from human- and lightning-caused fires in recent years (Bonnett 2020). We should mention that we did preliminary analysis of lightning clusters in BC for an earlier time period (1990–2004), and although we found similar clusters for the Montane Cordillera and Taiga Plains ecozones, the Vancouver Island cluster was not present. We suggest that this should be examined further to evaluate whether lightning has increased in the area over time, as regional drought has likely had the effect of lowering fuel moisture and increasing its receptivity to ignition. If such is the case, the Vancouver Island cluster may indicate a change in the local fire regime that may persist and grow larger in the future, which is why we raise this point.

Some of the more localised lightning-ignition clusters we identified using a lower hyperparameter setting seemed to be associated with roads. Previous research has shown an increased frequency of lightning fire with road density, possibly due to a higher proportion of flammable fuels along roadsides compared with forested areas (Arienti *et al.* 2009), which may help to explain this pattern. Some population centres also were within localised lighting fire clusters as well as human-caused fire clusters (e.g. Kamloops), which may highlight communities that area especially at risk from wildfire. One consideration is that the cause of wildfires may be misclassified, with human-caused fires attributed to lightning, which may have artificially elevated the number of lightning-caused fires near population centres and roads. Misclassification of wildfires maybe simply be due to human error; however, in some instances it may be due to the fact that it is easier to classify a fire as lightning-caused because it requires less effort (e.g. paperwork) than for human-caused fires.

Research has shown that spatial patterns in lightning-caused fires are related to lightning activity as well as various other factors (Wotton and Martell 2005). For example, lightning activity is highly influenced by proximity to cold water bodies, elevation and elevated terrain features (Orville *et al.* 2002; Burrows and Kochtubajda 2010). Furthermore, the type, amount, structure, continuity and moisture content of fuels are critical factors in the occurrence and spread of lightning-caused wildland fires (Wotton *et al*. 2003; Flannigan *et al.* 2016). Lightning-caused fires in the central Cordillera of western Canada, for instance, have been associated predominantly with coniferous forests (Wierzchowski *et al.* 2002).

Likewise, we found that both lightning- and human-caused ignitions in BC were mostly found in coniferous fuel types (Table 2), which make up the majority of fuels in the province (Supplementary Material S1, Supplementary Fig. S4). However, there were no human- or lightning-ignition cluster points found in the C3 fuels, which are primarily found in the northwestern part of the province. Another important point is that human-caused ignition points in clusters occurred in higher proportions in deciduous and mixed-wood fuel types than lightning-caused fires, which suggests that humans are an important factor in igniting fuels in these less flammable fuel types. Human-caused ignition points also occurred in higher proportions in the less represented C7 fuel type (Douglas fir and ponderosa pine), which is found in a part of the province with a fairly high level of human activity, including the Okanagan Valley, Kamloops and Merritt clusters. Lightning-ignition cluster points were also found in the highest proportions in the C5 and C2 fuel types. Similarly, the highest proportions of non-cluster lightning-fire points were found in C2, followed by C5 fuels. Human-caused ignition cluster points also differed from non-cluster points in that they were more often found in higher proportions in the D1 and M1 fuel types, which are primarily found the northeast part of the province where the Fort St John cluster was located (Supplementary Material S1, Supplementary Fig. S26). Non-cluster human-ignition points were found in very low proportions in the D1 and Boreal deciduous and conifer mixedwood with 50% conifer or more (M1) fuel types, which suggests that they are occurring primarily in the fairly concentrated area indicated in our clustering.

There are a number of factors that may have influenced our wildfire clustering results. Ecologically, larger fire events have the potential to influence cluster statistics, especially when considering number of fires, as it takes time for a forested area to regrow vegetation to be available to burn. This time-lag could account for a lack of fire activity in an area over a shorter time-period (Parks *et al.* 2018). Small fires may also be under-reported (Bridge *et al.* 2005; Hanes *et al.* 2019) including fire events that occurred in previously burned areas or in unburned islands within a larger burn. Policies and recording practices may also differ over time, and data quality can sometime be an issue. For example, two lightning-caused fire data points in our analysis fell outside provincial boundaries within the province of Alberta (Fig. 4). Furthermore, agency practices regarding documenting mutual aid fires near or within city or town limits are situation-dependent; such human-caused fires may not be recorded because they fall within town limits and are therefore not in designated forested management areas and are left out of the dataset. Furthermore, our exclusion of Parks Canada data may influence clustering patterns.

It is also important to note that our clustering study is primarily concerned with identifying regional clustering that offers visually apparent identification of spatial patterns and regions of possible interest for wildfire managers, researchers and other stakeholders. This work thus opens the door for future research aimed at understanding the drivers behind such clusters. Fortunately, previous research has examined some of the driving factors behind both lightning- and human-caused fires (as cited herein). For example, fuel moisture has been shown to play a critical role in the ignition of lightning- and human-caused wildland fires. Fires that occur after cloud-to-ground lightning strikes have been related to the moisture content of the organic layer of the forest floor, in particular, the layer represented by the Duff Moisture Code (DMC) in the Canadian Fire Weather Index (FWI) System, where fires can smoulder until surface fuels become dry enough to sustain surface spread (Wotton *et al*. 2003). In contrast to lightning-caused ignitions, human-caused fires have been shown to be more related to the moisture content of fine surface fuels, as represented by the Fine Fuel Moisture Code (FFMC) in the FWI System (e.g. twigs, needles and other cured fine fuels), at the time of contact with an ignition agent (e.g. firebrands from campfire embers or cigarettes; Cunningham and Martell 1973). Further work using machine learning clustering techniques, or other methods, may yield further insight into the factors influencing relationships between wildland fire, lightning and people. Further investigation is especially important because climate change is likely to lead to increasing fire weather extremes and increase the impact of wildland fire on the Canadian landscape. This may be especially important for the province of BC, which continues to experience noteworthy environmental challenges that may be exacerbated owing to climate change effects.

Conclusions

Here, we used the unsupervised machine learning algorithm HDBSCAN to characterise and compare high-density geographic clustering of human- and lightning-caused ignitions in western Canada. This method provided a robust and visually powerful approach for identifying areas in which both human- and lightning-caused wildfires occur in the highest densities and thus may be of benefit for wildfire management planning. Consistent with the literature, high-density clusters of human-caused ignitions were associated with areas of human activity such as roads and population centres. High-density clusters of lightning-caused ignitions also appeared to be associated with previously identified patterns of lightning activity. Clusters of human- and lightning-caused ignitions often overlapped, suggesting that these areas may be of special concern.

In the future, there are likely to be more wildland fires in Canada owing to climate change (Flannigan *et al.* 2009; Wang *et al.* 2015). Lightning is expected to increase with climate change (Romps *et al.* 2014), including in the Arctic (Chen *et al.* 2021). Thus, spatial lightning patterns will likely change in the future, as they have in the past. Furthermore, future climatically driven increases in precipitation in some regions may not be enough to offset the effect of increased in temperature on fuel moisture, thereby leading to fuels that are more receptive to ignition (Flannigan *et al.* 2016). There will also likely be more interface areas in the future (Theobald and Romme 2007). Therefore, the above-mentioned factors, in combination with the future challenges that are expected for fire management and suppression (Podur and Wotton 2010), make it likely that there will be more wildland fire in Canada in the future, including an increase in destructive interface fire events. Machine learning and other analytical approaches are likely to play an increasing role in the management of wildland fire in Canada, as fire regimes continue to change and wildland fire is likely to continue to become increasingly challenging to manage.

Supplementary material

Supplementary material is available online.

References

Arienti MC, Cumming SG, Krawchuk MA, Boutin S (2009) Road network density correlated with increased lightning fire incidence in the Canadian western boreal forest. *International Journal of Wildland Fire* **18**, 970–982. doi:10.1071/WF08011

Balch JK, Bradley BA, Abatzoglou JT, Nagy RC, Fusco EJ, Mahood AL (2017) Human-started wildfires expand the fire niche across the United States. *Proceedings of the National Academy of Sciences* **114**(11), 2946–2951. doi:10.1073/pnas.1617394114

Bistinas I, Oom D, Sá ACL, Harrison SP, Prentice IC, Pereira JMC (2013) Relationships between human population density and burned area at continental and global scales. *PLoS One* **8**(12), e81188. doi:10.1371/journal.pone.0081188

Bonnett N (2020) Climate change adaptations planning at the regional district scale on Vancouver Island, British Columbia. Master's thesis, University of Alberta, Edmonton, Canada. doi:10.7939/r3-gjyc-0z35

Bridge SRJ, Miyanishi K, Johnson EA (2005) A critical evaluation of fire suppression effects in the boreal forest of Ontario. *Forest Science* **5**, 41–50. doi:10.1093/forestscience/51.1.41

Burrows WR, Kochtubajda B (2010) A decade of cloud-to-ground lightning in Canada: 1999-2008. Part 1: Flash density and occurrence. *Atmosphere–Ocean* **48**(3), 177–194. doi:10.3137/AO1118.2010

Campello RJGB, Moulavi D, Sander J (2013) Density-based clustering based on hierarchical density estimates. In 'Advances in knowledge discovery and data mining. PAKDD 2013. Lecture notes in computer science. vol. 7819'. (Eds J Pei, VS Tseng, L Cao, H Motoda, G Xu) pp. 160–172. (Springer: Berlin, Heidelberg) doi:10.1007/978-3-642-37456-2_14

Canadian Forest Service (2021) 'Canadian National Fire Database – Agency Fire Data.' (Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre: Edmonton, Alberta) Available at http://cwfis.cfs.nrcan.gc.ca/datamart

Caswell TA, Droettboom M, Lee A, Sales de Andrade E, Hunter J, Hoffmann T, *et al.* (2021) 'matplotlib/matplotlib: REL: v3.4.0 (Version v3.4.0).' (Zenodo) doi:10.5281/zenodo.4638398

Chen Y, Romps DM, Seeley JT, Veraverbeke S, Riley WJ, Mekonnen ZA, Randerson JT (2021) Future increases in Arctic lightning and fire risk for permafrost carbon. *Nature Climate Change* **11**, 404–410. doi:10.1038/s41558-021-01011-y

Coogan SCP, Robinne F-N, Jain P, Flannigan MD (2019) Scientists' warning on wildfire — a Canadian perspective. *Canadian Journal of Forest Research* **49**, 1015–1023. doi:10.1139/cjfr-2019-0094

Coogan SCP, Cai X, Jain P, Flannigan MD (2020) Seasonality and trends in human- and lightning-caused wildfires ≥2 ha in Canada, 1959–2018. *International Journal of Wildland Fire* **29**, 473–485. doi:10.1071/WF19129

Coogan SCP, Daniels LD, Boychuk D, Burton PJ, Flannigan MD, Gauthier S, Kafka V, Park JS, Wotton BM (2021) Fifty years of wildland fire science in Canada. *Canadian Journal of Forest Research* **51**, 283–302. doi:10.1139/cjfr-2020-0314

Cunningham AA, Martell DL (1973) A stochastic model for the occurrence of man-caused forest fires. *Canadian Journal of Forest Research* **3**, 282–287. doi:https://doi.org/10.1139/x73-038

Ester M, Kriegel H-P, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In 'Proceedings of the Second International Conference on Knowledge Discovery and Data Mining'. (Eds E Simoudis, J Han, U Fayyad) pp. 226–231. (AAAI Press)

Flannigan MD, Krawchuk MA, de Groot WJ, Wotton BM, Gowman LM (2009) Implications of changing climate for global wildland fire. *International Journal of Wildland Fire* **18**, 483–507. doi:10.1071/WF08187

Flannigan MD, Wotton BM, Marshall GA, de Groot WJ, Johnston J, Jurko N, Cantin AS (2016) Fuel moisture sensitivity to temperature and precipitation: climate change implications. *Climatic Change* **134**, 59–71. doi:10.1007/s10584-015-1521-0

Government of BC (2021) Wildfire season summary. Available at https://www2.gov.bc.ca/gov/content/safety/wildfire-status/about-bcws/wildfire-history/wildfire-season-summary

Gralewicz NJ, Nelson TA, Wulder MA (2012) Factors influencing national scale wildfire susceptibility in Canada. *Forest Ecology and Management* **265**, 20–29. doi:10.1016/J.FORECO.2011.10.031

Hanes CC, Wang X, Jain P, Parisien M-A, Little JM, Flannigan MD (2019) Fire-regime changes in Canada over the last half century. *Canadian Journal of Forest Research* **49**, 256–269. doi:10.1139/CJFR-2018-0293

Hennig C (2015) What are the true clusters? *Pattern Recognition Letters* **64**, 53–62. doi:10.1016/j.patrec.2015.04.009

Jain P, Coogan SCP, Subramanian SG, Crowley M, Taylor S, Flannigan MD (2020) A review of machine learning applications in wildfire science and management. *Environmental Reviews* **28**, 478–505. doi:10.1139/er-2020-0019

Jain P, Castellanos-Acuna D, Coogan SCP, Abatzoglou JT, Flannigan MD (2021) Observed increases in extreme fire weather driven by atmospheric humidity and temperature. *Nature Climate Change* **12**, 63–70. doi:10.1038/s41558-021-01224-1

Johnston LM, Flannigan MD (2018) Mapping Canadian wildland fire interface areas. *International Journal of Wildland Fire* **27**, 1–14. doi:10.1071/WF16221

Krawchuk MA, Cumming SG, Flannigan MD, Wein RW (2006) Biotic and abiotic regulation of lightning fire initiation in the mixed-wood boreal forest. *Ecology* **87**(2), 458–468. doi:10.1890/05-1021

Lewis J (2009) A psychophysical investigation of clustering. In 'Proceedings of the 34th Annual Conference of the Cognitive Science Society, Austin, Texas, 2012'. (Eds NA Taatgen, H van Rijn) pp. 1870–1875. (Cognitive Science Society)

Maria E, Budiman E, Haviluddin H, Taruk M (2020) Measure distance locating nearest public facilities using Haversine and Euclidean methods. *Journal of Physics: Conference Series* **1450**, 012080. doi:10.1088/1742-6596/1450/1/012080

McGee TK (2021) Evacuating First Nations during wildfires in Canada. *Fire Safety Journal* **120**, 103120. doi:10.1016/j.firesaf.2020.103120

McInnes L, Healy J, Astels S (2017) hdbscan: Hierarchical density based clustering. *The Journal of Open Source Software* **2**(11), 205. doi:10.21105/joss.00205

Meyn A, Schmidtlein S, Taylor SW, Girardin MP, Thonicke K, Cramer W (2010) Spatial variation of trends in wildfire and summer drought in British Columbia, Canada, 1920–2000. *International Journal of Wildland Fire* **19**, 272–283. doi:10.1071/WF09055

Moore RD, Spittlehouse DL, Whitfield PH, Stahl K (2010) Weather and climate. In 'Compendium of forest hydrology and geomorphology in BC'. (Eds RG Pike, TE Redding, RD Moore, RD Winkler, KD Bladon) pp. 47–84. (BC Ministry of Forests and Range, Research Branch and FORREX Forest Research Extension Partnership: Victoria and Kamloops, BC, Canada)

Moulavi D, Jaskowiak PA, Campello RJGB, Zimek A, Sander J (2014) Density-based clustering validation. In 'Proceedings of the 2014 SIAM International Conference on data mining (SDM)'. pp. 839–847. (SIAM) doi:10.1137/1.9781611973440.96

Nadeem K, Taylor SW, Woolford DG, Dean CB (2020) Mesoscale spatiotemporal predictive models of daily human- and lightning-caused wildland fire occurrence in British Columbia. *International Journal of Wildland Fire* **29**, 11–27. doi:10.1071/WF19058

Orville RE, Huffines GR, Burrows WR, Holle RL, Cummins KL (2002) The North American lightning detection network (NALDN)—first results: 1998–2000. *Monthly Weather Review* **130**, 2098–2109. doi:10.1175/1520-0493(2002)130<2098:TNALDN>2.0.CO;2

Parks SA, Parisien MA, Miller C, Holsinger LM, Baggett LS (2018) Fine-scale spatial climate variation and drought mediate the likelihood of reburning. *Ecological Applications* **28**, 573–586. doi:10.1002/eap.1671

Petoukhov V, Petri S, Kornhuber K, Thonicke K, Coumou D, Schellnhuber HJ (2018) Alberta wildfire 2016: apt contribution from anomalous planetary wave dynamics. *Scientific Reports* **8**, 12375. doi:10.1038/s41598-018-30812-z

Pew KL, Larsen CPS (2001) GIS analysis of spatial and temporal patterns of human-caused wildfires in the temperate rain forest of Vancouver Island, Canada. *Forest Ecology and Management* **140**, 1–18. doi:10.1016/S0378-1127(00)00271-1

Podur J, Wotton M (2010) Will climate change overwhelm fire management capacity? *Ecological Modelling* **221**, 1301–1309. doi:10.1016/j.ecolmodel.2010.01.013

Price O, Bradstock R (2014) Countervailing effects of urbanization and vegetation extent on fire frequency on the wildland–urban interface: disentangling fuel and ignition effects. *Landscape and Urban Planning* **130**, 81–88. doi:10.1016/J.LANDURBPLAN.2014.06.013

Robinne F-N, Parisien M-A, Flannigan M (2016) Anthropogenic influence on wildfire activity in Alberta, Canada. *International Journal of Wildland Fire* **25**, 1131–1143. doi:10.1071/WF16058

Robusto CC (1957) The cosine-haversine formula. *The American Mathematical Monthly* **64**, 38–40. doi:10.2307/2309088

Romps DM, Seeley JT, Vollaro D, Molinari J (2014) Projected increase in lightning strikes in the United States due to global warming. *Science* **346**, 851–854. doi:10.1126/science.1259100

Schiermeier Q (2021) Climate change made North America's deadly heatwave 150 times more likely. *Nature.* 8 July 2021. Available at https://www.nature.com/articles/d41586-021-01869-0

Stocks BJ, Mason JA, Todd JB, Bosch EM, Wotton BM, Amiro BD, Flannigan MD, Hirsch KG, Logan KA, Martell DL, Skinner WR (2002) Large forest fires in Canada, 1959–1997. *Journal of Geophysical Research: Atmospheres* **107**(D1), 8149. doi:10.1029/2001JD000484

Theobald DM, Romme WH (2007) Expansion of the US wildland–urban interface. *Landscape and Urban Planning* **83**, 340–354. doi:10.1016/J.LANDURBPLAN.2007.06.002

Tymstra C, Stocks BJ, Cai X, Flannigan MD (2020) Wildfire management in Canada: review, challenges and opportunities. *Progress in Disaster Science* **5**, 100045. doi:10.1016/j.pdisas.2019.100045

USDA, USDI (2001) 'Notices'. pp. 751–777. (Federal Register, United States Government) Available at https://www.federalregister.gov/articles/2001/01/04/01-52/urban-wildland-interface-communities-within-the-vicinityof-federal-lands-that-are-at-high-risk-from [accessed 18 February 2021]

Wang X, Thompson DK, Marshall GA, Tymstra C, Carr R, Flannigan MD (2015) Increasing frequency of extreme fire weather in Canada with climate change. *Climatic Change* **130**, 573–586. doi:10.1007/S10584-015-1375-5

Wang X, Parisien M-A, Taylor SW, Candau J-N, Stralberg D, Marshall GA, Little JM, Flannigan MD (2017) Projected changes in daily fire spread across Canada over the next century. *Environmental Research Letters* **12**, 025005. doi:10.1088/1748-9326/aa5835

Wierzchowski J, Heathcott M, Flannigan MD (2002) Lightning and lightning fire, Central Cordillera, Canada. *International Journal of Wildland Fire* **11**, 41–51. doi:10.1071/WF01048

Wotton BM, Martell DL (2005) A lightning fire occurrence model for Ontario. *Canadian Journal of Forest Research* **35**, 1389–1401. doi:10.1139/X05-071

Wotton BM, Martell DL, Logan KA (2003) Climate change and people-caused forest fire occurrence in Ontario. *Climatic Change* **60**, 275–295. doi:10.1023/A:1026075919710

Wotton BM, Flannigan MD, Marshall GA (2017) Potential climate change impacts on fire intensity and key wildfire suppression thresholds in Canada. *Environmental Research Letters* **12**, 095003. doi:10.1088/1748-9326/aa7e6e

Data availability. Data used in this study are available from Natural Resources Canada online at https://cwfis.cfs.nrcan.gc.ca/datamart.

Conflicts of interest. Coauthor Mike D. Flannigan is an Associate Editor for the International Journal of Wildland Fire. He had no editorial involvement with this paper, which was handled independently by other members of the Editorial Board. The authors declare no other conflicts of interest.

Declaration of funding. Funding provided by Canada Wildfire Natural Sciences and Engineering Research Council of Canada (NSERC) Strategic Network.

Acknowledgements. We thank Ginny Marshall for her assistance with the aggregated fuels raster.

Author contributions. SCPC conceived the study, analysed the data and wrote the first draft of the manuscript. OA, MF and SCPC reviewed and revised the manuscript. All authors contributed intellectually to the manuscript.

Author affiliations

[A]Department of Renewable Resources, University of Alberta, Edmonton, AB T6G 2H1, Canada.

[B]Department of Natural Resource Science, Thompson Rivers University, Kamloops, BC V2C 0C8, Canada.