

Predicting Wildfire Occurrence in British Columbia with a Spatio-Temporal Graph Neural Network and Multi-Source Environmental Data

Shiheng Huang

Department of Statistical Science
University of Toronto

Abstract

Wildfire prediction has become increasingly critical as fire activity intensifies under climate change, yet traditional systems often rely on interpolated weather indices. This study develops a Spatio-Temporal Graph Neural Network (ST-GNN) framework to model the complex spatiotemporal dependencies for next-day wildfire occurrence prediction in British Columbia, Canada. We integrated and aligned historical fire records from 2019 to 2024 with corresponding ERA5 reanalysis variables, terrain features, MODIS NDVI indices, land cover data, and a suite of engineered features to a 0.25-degree spatial grid. Next, we constructed a spatial graph to encode geographical proximity using K-nearest neighbours (KNN), and utilized a bidirectional Long Short-Term Memory (LSTM) to capture temporal dynamics. The model was specifically trained with focal loss to address the extreme class imbalance in the data, and evaluated under five model configurations. The results reveal that the ST-GNN model successfully identifies the spatiotemporal patterns of wildfire risk across all configurations. The inclusion of engineered features enhanced predictive capability, while shorter input sequence length (7-day) increased model sensitivity and longer windows (21-day) improved precision. The model performs strongly in peak fire season (April to September) with a recall peaking at 0.47 in July, but its effectiveness declines during the low-activity season. These findings demonstrate that explicitly modelling spatiotemporal dependencies can supplement existing wildfire risk systems and offer a scalable approach for early wildfire warning.

Keywords: Wildfire Prediction, Graph Neural Networks, Spatiotemporal Modelling, Multi-source Environmental Data, British Columbia

Contents

1	Introduction	4
2	Data	6
2.1	BC Boundary	6
2.2	Fire Points	6
2.3	ERA5	10
2.4	NDVI	14
2.5	DEM	15
2.6	Landcover	16
3	Feature Engineering	16
3.1	Fire Labels	17
3.2	ERA5 Feature Selection	17
3.3	Bad Grids	18
3.4	Feature Merge	18
3.5	Features Addition	18
4	Methodology	20
4.1	Datasets Partition	20
4.2	Spatial Graph	20
4.3	ST-GNN Architecture	21
4.3.1	Input Embedding Layer	22
4.3.2	Spatial Graph Convolution Layer	22
4.3.3	Temporal Sequence Modelling Layer	23
4.3.4	Prediction Output Layer	23
4.4	Loss Function	24
5	Results	24
5.1	Threshold Optimization	25
5.2	Basic Classification	26
5.3	Temporal Analysis	27
5.4	Spatial Analysis	28
6	Discussion	30
6.1	Software availability	35
6.2	Tables and Figures	36

1 Introduction

Wildfire has transformed from a natural ecological process into an environmental hazard over the past decades around the globe. Human-caused climate change has been widely attributed to this shift toward rising temperatures and prolonged drought, creating conducive conditions for extreme fire behaviour, and resulting in an increasing frequency and intensity of wildfires (Abatzoglou & Williams, 2016). Canada exemplifies this escalating crisis, experiencing its worst wildfire season on record in 2023, with over 17 million hectares burned and hundreds of thousands of residents displaced (Canadian Interagency Forest Fire Centre, 2023). This devastated the economy and the human community, while smoke from these fires affected air quality across North America.

British Columbia serves as an important case study within this global trend. The province's topography ranges from coastal temperate rainforests through multiple parallel mountain systems to inland plateaus and alpine zones from west to east, spanning an elevation of 4663 meters above sea level (Church & Ryder, 2010). This pronounced heterogeneous landscape creates a complex and spatially variable wildfire regime.

Historical records indicate that BC experiences an average of 1,500 wildfires annually, predominantly occurring between April and September. While this seasonal pattern has persisted, the annual area burned has increased from a 25-year average of 383,853 hectares to a five-year average of 982,719 hectares, with at least 800,000 hectares burned annually since 2021 (Canadian Interagency Forest Fire Centre, 2025). The 2017, 2018, and 2023 fire seasons each saw unprecedented fire activity levels, culminating in 2023's record-breaking season with 2.84 million hectares burned and 481 simultaneous active fires during the peak season (BC Wildfire Service, 2025).

The existing wildfire risk management framework in BC operates within the nationally standardized Canadian Forest Fire Danger Rating System (CFFDRS). One core component Fire Weather Index (FWI) employs a semi-empirical physical model to convert weather station observations of temperature, relative humidity, wind speed, and precipitation into numerical indices that represent fuel moisture conditions and fire danger levels (Wagner, 1987). The outputs of the FWI system are subsequently combined with information on fuel type and terrain slope within the Fire Behaviour Prediction (FBP) system, which quantitatively estimates fire behaviour such as fire spread rates and intensity (Wotton, 2009).

Although the CFFDRS is mature and widely implemented in practical fire management, it has several limitations for landscape-scale wildfire risk prediction. First, the FWI computation heavily relies on a sparse weather station network, resulting in spatially interpolating the fire hazard index rather than the actual observations. This interpolation may fail to

accurately represent the condition between observation points to alert in advance, particularly in topographically complex regions. On the other hand, the system operates as a deterministic model using daily meteorological inputs, which treats each location and time step independently. This approach lacks mechanisms to capture complex spatiotemporal dependencies, such as spatial fire spread from neighbouring areas or the temporal persistence effect. In addition, the warning thresholds within the system are typically derived from the historical statistical distribution. However, as climate change drives increasingly frequent extreme fire weather events and shifts fire patterns, the reliability of the threshold-based system may be diminishing.

To address these limitations, we propose a spatiotemporal modelling framework based on Graph Neural Networks (GNNs) (Scarselli et al., 2008) to capture complex spatiotemporal patterns in historical wildfire observations. Graph Neural Networks explicitly represent spatial relationships through graph structures, in which geographic units are modelled as nodes connected by edges that encode spatial proximity. Temporal convolution mechanisms are further integrated to jointly model spatial dependencies and temporal dynamics, thereby addressing the spatiotemporal modelling gap inherent in current operational wildfire risk assessment systems (Yu et al., 2018).

This thesis has three primary objectives. First, we characterize the spatiotemporal pattern of BC wildfires during 2019-2024 using exploratory spatial statistics and correlation analysis. Second, we develop a feature engineering framework that integrates weather analysis data, terrain characteristics, vegetation indices, and land cover information into a spatiotemporal structured predictor, which is suitable for GNN architectures. Third, we design and evaluate an ST-GNN model for next-day wildfire occurrence prediction, and compare the performance of different predictors as input. Through these analyses, we aim to demonstrate whether explicitly modelling spatiotemporal dependencies can improve the accuracy in wildfire prediction, and provide interpretable insights for operational fire management.

This paper is organized as follows. In Section 2, we provide a detailed description of the sample datasets, including data manipulation and exploratory analyses of each dataset. Next, we outline the feature engineering strategies we incorporated into the datasets in Section 3. Then, in Section 4, we present the proposed spatiotemporal Graph Neural Network (ST-GNN) framework to estimate the probability of wildfire occurrence at specific locations and dates, and we illustrate the corresponding results in Section 5. Finally, the discussion section summarizes the main findings of the study and provides implications and future research directions.

2 Data

We assembled six spatiotemporal datasets to address the challenge of capturing wildfire occurrence predictions, including fire points patterns, weather conditions, terrain characteristics, and vegetation state. These datasets encompass six complete fire seasons from 2019 to 2024, including both typical and extreme fire years. To balance the need for spatial detail with computational feasibility, we partitioned BC into a regular grid of 2,367 analysis units at 0.25-degree resolution (approximately 25 km x 25 km per grid cell). Each grid cell serves as a node in our graph structure, while daily observations form the temporal dimension. All spatial datasets were standardized to the EPSG:4326 coordinate reference system (WGS84) prior to analysis.

2.1 BC Boundary

The administrative boundary of British Columbia was obtained from the 2021 Census Boundary Files provided by Statistics Canada (Statistics Canada, 2021). This dataset delineates the official extent of British Columbia at the provincial level and was used to define the spatial domain of the study. The study area was partitioned into 2,367 uniform grids, which were used to mask and spatially align all datasets to a consistent geographic extent. The boundary geometry was also processed and projected to EPSG:4326 to ensure spatial consistency across data layers.

2.2 Fire Points

Historical wildfire records were obtained from the Canadian Wildland Fire Information System (Canadian Forest Service, 2024). The dataset contains point records of all reported fire incidents from 2019 to 2024 in BC, comprises 8,939 fire points in total. Each fire record includes the incident location, reported date, put out date, fire cause classification and burned area, and all variables are listed in Table S1.

The baseline years of 2019 and 2020 recorded relatively modest wildfire activity, totalling 1,502 fires across the two years as shown in Figure 1A. However, 2021 marked a pivotal transition, with a sharp increase to 1,649 fires. This upward trend persisted through 2022 and culminated in 2023, which experienced the highest fire frequency on record with 2,280 incidents. Although fire counts declined slightly in 2024 to 1,725, the post-2020 period averaged 1,859 fires per year is still approximately 2.5 times the 2019 and 2020 baseline. Collectively, these patterns suggest a sustained shift in wildfire activity rather than short-lived stochastic fluctuations, and potentially reflecting the influence of persistent climate

change.

Figure 1B illustrate that wildfire activity exhibits significant seasonal concentration, with April to September consistently accounting for 94.3% of all fires across the six-year study period. Within this season, July and August represent the peak months. July alone recorded 3,070 fires, while August accounted for 2,535 fires, together comprising 62.6% of all wildfire occurrences during the study period.

Critically, the divergence in coefficients of variation between fire counts and burned area demonstrated that frequent fires are not the main consequence of extensive burned area. This distinction is clearly shown by contrasting within the study period. Despite the year 2022 recording the second-highest fire counts of 1,783, the average burned area is only 76 hectares per incident in Figure S1A. While this yearly average is still three times the baseline periods, but extremely low when compared to the extensive fire years of 2021 (525 ha), 2023 (627 ha), and 2024 (1,245 ha) per incident.

This deviation may be derived from the delayed or extended fire season in 2022, as shown in Figure S1B. Unlike the three-year average for extensive fire years, where July accounted for 45% of annual fire, 2022 has only recorded 15.1% of annual fire with 269 fires in July. Fire activity instead peaked in August, which contained 51.4% (917) of annual fires compared to the three extensive year average of 21.5% for August. Additionally, 2022 experienced an unusually prolonged fire season, with 166 fires occurring in October, while the six-year average in October is 48 fires per year. This temporal displacement likely reflected anomalous weather conditions in 2022, potentially including delayed drought conditions, cooler temperatures, or increased precipitation frequency during the peak fire month. Most importantly, this meteorological difference may have also facilitated more effective initial attack and containment during the fire management to lower the burned area with high fire count.

In contrast to summer intensity, fire activity is nearly absent during the November to March, with only 216 fires recorded over 6 years. January and February each recorded only 10 fires across the entire study period, demonstrating that the winter conditions effectively suppress both fire ignition and spread.

Beyond temporal patterns, spatial distribution is also critical for understanding fire occurrence and spread dynamics. To investigate spatial structure, we overlaid the 8,939 fire incidents from 2019 to 2024 onto the 2,367 analysis grids covering British Columbia, and the structure is shown in Section 3.1. Of these grids, 1,373 grids experienced at least one fire during the study period, while the remaining 995 remained fire-free. The fire count distribution strongly demonstrates a positive skew and heavy tail, as the median grid recorded only 1 incident over six years while the 75th percentile reached 4 fires. In contrast, the maximum

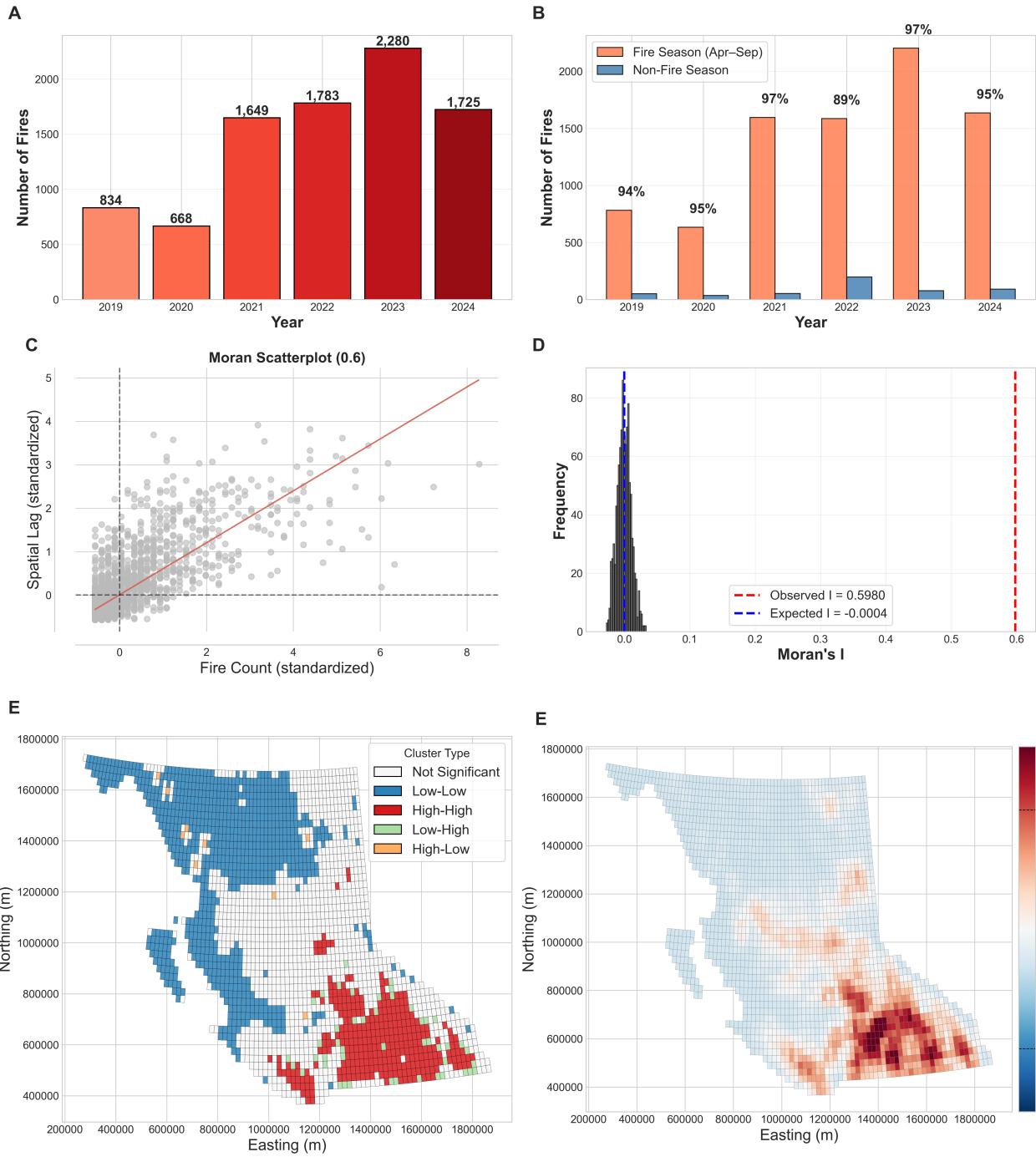


Figure 1: (A) Annual wildfire counts in British Columbia from 2019–2024; (B) Comparison of fire-season (April–September) and non-fire-season fire counts by year; (C) Global Moran’s I scatterplot of grid-level fire counts ($I = 0.598$, $p = 0.001$); (D) Permutation distribution of Moran’s I under spatial randomization; (E) Local Indicators of Spatial Association (LISA) cluster map; (F) Getis–Ord Gi^* z-scores identifying local hot and cold spots.

number of fires within a single grid reached 58, whereas the mean fire count was 3.78 with a standard deviation of 6.67. This overdispersion indicates that fire risk is highly spatially concentrated rather than uniformly distributed.

The annual spatial comparison in Figure S1C displays the fire locations between 2021 to 2023 for each year. These figures reveal that the spatial stability is consistent despite variation in total fire counts. Across all years, the southeastern region (Easting: 1,200,000 - 1,600,000m, Northing: 400,000 - 800,000 m) consistently exhibits the highest fire density, with numerous grids recording 5 or more fires annually in extreme years. While 2021 showed moderate concentration in the southeastern cluster with scattered activity in Central BC, wildfire occurrence in 2022 became more spatially dispersed across the province. In contrast, 2023 demonstrated the most intense clustering with dark red cells, and subsequently expanded into the typically moderate risk regions in the northward direction. The spatial extent of these hotspots partially contracts in 2024, although core high activity remains persistent.

We further provided strong statistical evidence that fire occurrence is highly non-random across space. The observed Global Moran's I of 0.598 substantially exceeds the expected value under spatial randomness, given by $\frac{-1}{(n-1)}$, equals -0.004 (Moran, 1950). This confirms strong positive spatial autocorrelation at the 0.25-degree grid scale. The Moran scatterplot in Figures 1C demonstrated observations fall sparsely Quadrant I (High-High), representing grids with above-average fire counts surrounded by similarly high fire count neighbours (Rey et al., 2019). The majority of grids cluster in Quadrant III (Low-Low), corresponding to coastal and northern regions with low fire counts surrounded by low-fire neighbours. The remainder proportion of grids are located at the spatial outlier quadrants (High-low and Low-High). These grids represent isolated high-fire grids within low-fire regions or vice versa, indicating these regions are uncommon and that fire risk transitions gradually rather than with sharp discontinuities.

The Local Indicators of Spatial Association (LISA) in Figure 1E identifies 1,244 grids with statistically significant local spatial autocorrelation at the 5% significance level (Anselin, 1995). The High-High clusters ($n=322$) represent the persistent fire hotspot core in the southeastern interior, where these grids likely share climate characteristics and vegetation composition that systematically elevate fire risk. The coastal BC and far northern regions contain the majority of the Low-Low clusters ($n=861$), forming a naturally consistent fire-free region. The Low-High outliers ($n=52$) are potentially reflecting recent burn scars or protected areas, while the High-Low outliers ($n=9$) may reflect localized human influence or recent anomalous fire activity.. Complementary Getis-Ord Gi* analysis in Figure 1F identifies 94 significant hotspot grids (Getis & Ord, 1992). At the 95% confidence level, 65

grid cells are identified as robust wildfire hotspots, while only 29 cells remain significant at the 99% level. These grids overlapped with LISA clusters in Figure 1E, confirming the spatial consistency between the two methods and the existence of spatial autocorrelation in fire locations.

2.3 ERA5

Weather variables were obtained from the ERA5 global atmospheric reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) and accessed via Google Earth Engine (Hersbach et al., 2020; Gorelick et al., 2017). The dataset provides gridded, physically consistent estimates of near-surface meteorological conditions. In this study, daily weather variables from Jan 1, 2019 to Dec 31, 2024 (2,192 days in total) covering the province of British Columbia were extracted, following the official province extent from Statistics Canada mentioned in Section 2.1. The datasets include comprehensive measures of air temperature, precipitation, wind, radiation, soil moisture, and snow-related variables. A total of 23 covariates were extracted, and detailed variable descriptions are provided in Table S2.

We selected key meteorological variables from the ERA5 dataset to conduct exploratory data analysis relevant to wildfire activity in British Columbia. Figure 2A shows the 7-day rolling mean of daily temperature aggregated across the province, which suppresses short-term variability and highlights pronounced seasonal patterns. The resulting time series presents a clear annual cycle, with summer maxima occurring between July and August, while winter minima occur in January and February. Daily mean temperatures across BC show an annual range of 30 to 40 degrees Celsius, where rapid spring warming transitions between March and May, and more gradual autumn cooling from September through November.

Substantial interannual deviations from the six-year mean are evident during the study period. In particular, the summer of 2021 exhibited exceptional warmth in the summer, with peak temperatures exceeding 20 degrees Celsius and sustained above-average temperatures throughout July and August. This period coincided with the severe wildfire season in 2021, which recorded 1,649 fires and 866 thousand hectares burned. Similarly, 2023 displayed persistently high fire season temperatures, with July mean temperatures approaching 19 degrees Celsius. These extended warm periods aligned with the frequent fire occurrence rates in 2023. In contrast, 2020 and 2022 exhibited more moderate summer temperatures closer to the multi-year average. Specifically in 2022, the temperatures remained below average when the fire season commenced in April, peaking above average in August and

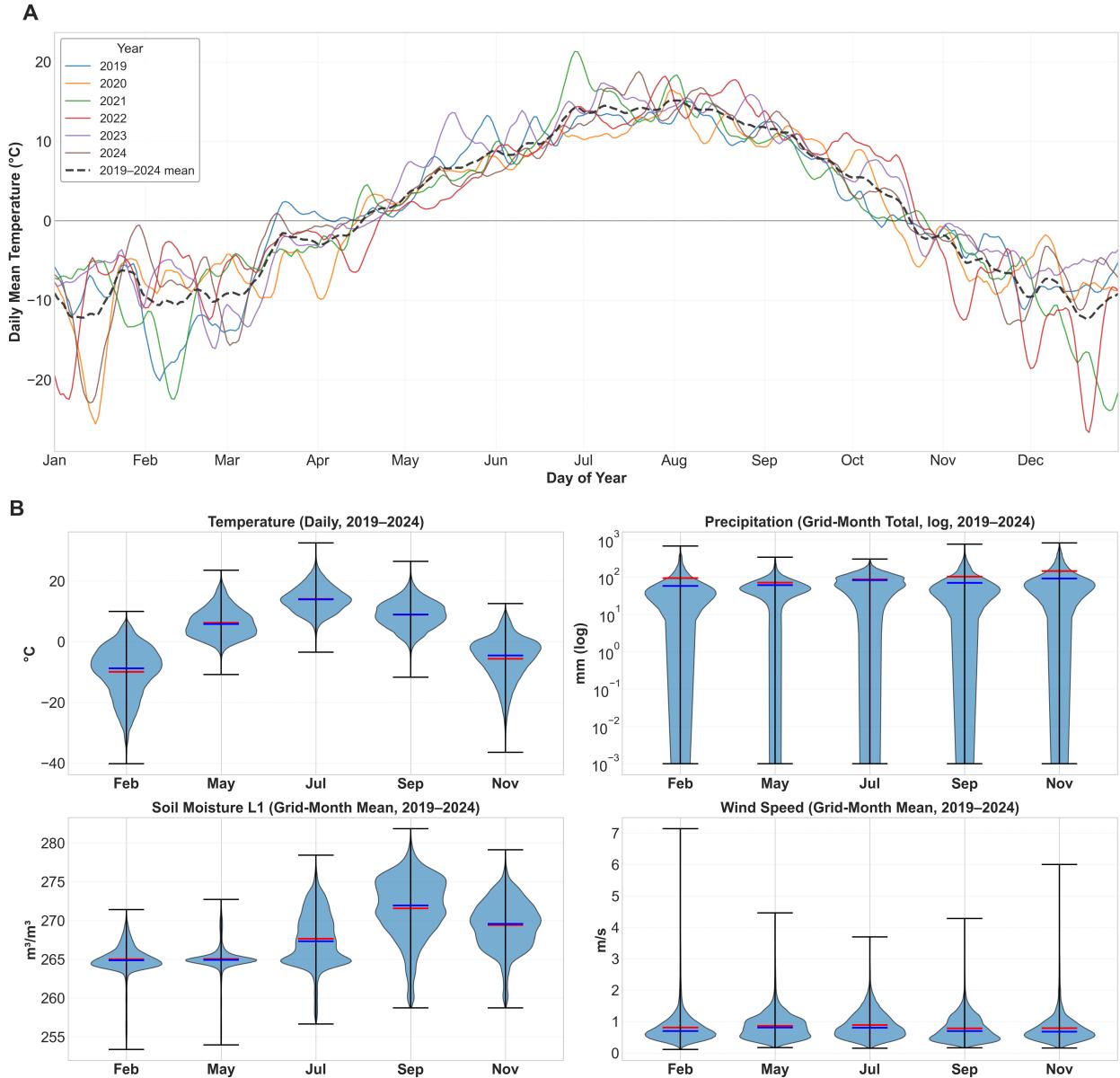


Figure 2: (A) Interannual comparison of daily mean temperature (7-day rolling mean) from 2019–2024; (B) Distributions of key weather variables for selected months over the same period, mean is coloured in red while median is coloured in blue.

remaining elevated through October. This temperature pattern corresponded closely with the delayed and extended fire season observed in 2022, as discussed in Section 2.2. Winter temperatures display comparatively low interannual variability, with all six years exhibiting similar cold-season conditions.

The violin plots in Figure 2B illustrate the monthly distributions of weather parameters across all years. These months were chosen to represent distinct seasonal regimens in BC. Temperature distributions demonstrate the expected seasonal progression. The transitional months of February and November show the broadest distributions, ranging from approximately -40 to -15 degrees Celsius. These wide ranges reflect the spatial heterogeneity of coastal areas remain temperate, while interior and northern regions experience early or late freezes. In contrast, the summer month of July displays the narrowest temperature distribution, indicating more spatially homogeneous warming conditions across the province. May and September show intermediate variability, with slightly greater spread likely attributable to differential spring warming rates across elevation gradients.

Grid-month total precipitation reveals distinct seasonal moisture patterns for each grid. While the median remains relatively consistent across seasons, extreme precipitation events occur throughout the year. February, September, and November exhibit the greatest variability, spanning nearly two orders of magnitude. July precipitation shows a slight bimodal pattern, potentially distinguishing between coastal regions with slightly higher values and the interior with lower rain shadow.

Soil moisture demonstrates strong seasonal dependence, with minimum values in July and peak values in September, consistent with enhanced summer evapotranspiration followed by reduced evaporative demand in autumn. Wind speed distribution reveals modest seasonal variability, with median values ranging from 0.6 to 1.0 m/s across all examined months. May and September tend to exhibit fuller distribution with extended tails, while July shows a higher proportion of wind speeds exceeding 1.5 m/s, potentially associated with convective activity during the fire season.

Spatial distributions of meteorological variables and wildfire occurrences for July 2024 are shown in Figure 3. The mean temperature in this peak fire month exhibits a clear Southeast to Northwest gradient. While the warmest regions are coloured in dark red, concentrated in the southern interior, the coolest temperatures are coloured in light blue, appearing across northern regions in BC. Overlaying fire incidents onto this temperature field reveals that fires are preferentially distributed in the warmer southeastern and central interior regions, which are concentrated in areas with moderate temperatures between 10 to 20 degrees Celsius. Despite moderate temperatures, coastal areas show lower fire density, likely attributed to concurrently high precipitation and moisture availability, as demonstrated in the subsequent

panels.

Total precipitation patterns display an inverse relationship with fire occurrence. The wettest regions, coloured in deep blue, align with the coastal mountains and western slopes and possibly orographic enhancement of Pacific moisture. These areas contain very few fire points, demonstrating that moisture can effectively limit fire activity. Notably, these high rain shadow zones also correspond to regions with the lowest temperatures yet relatively dry soil moisture values. Conversely, interior regions receiving moderate to low precipitation display more fire activity, particularly in areas that also experience higher temperatures.

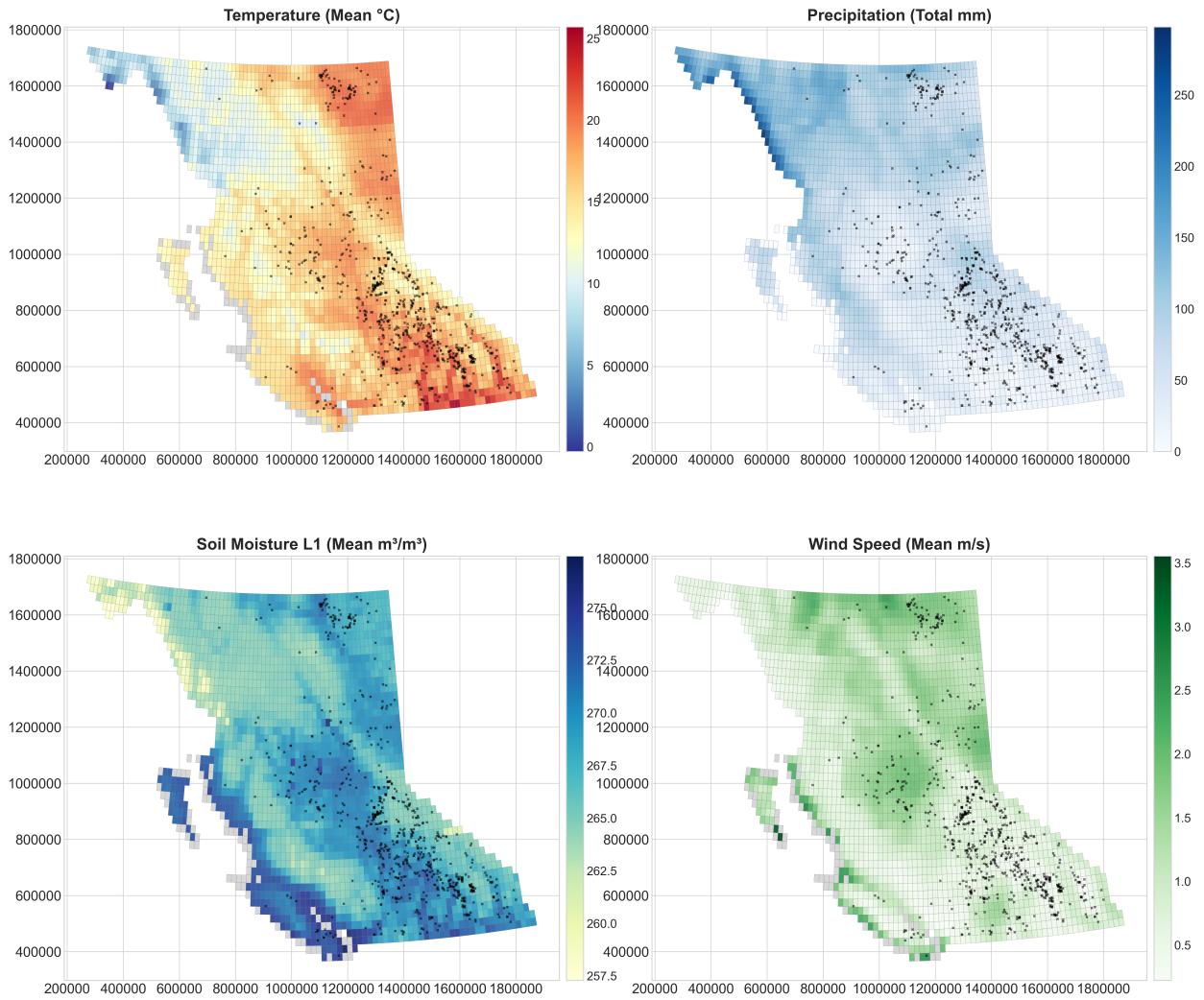


Figure 3: Spatial patterns of selected July 2024 weather variables across British Columbia, shown at the grid level with wildfire incident points overlaid.

The soil moisture spatial patterns diverge notably from precipitation distributions, where regions with relatively higher precipitation along the coastal slopes exhibit drier soil mois-

ture values. In contrast, the southern interior shows moderate soil moisture in light blue despite lower precipitation, potentially reflecting cumulative moisture retention from previous months. Counter-intuitively, these regions of moderate soil moisture present heavy overlay with peak fire occurrence, and are worth further investigation into the details. The southwestern coastal region maintains the highest soil moisture content with minimum precipitation, which is possibly influenced by the Pacific Ocean.

Mean wind speeds range from 0.5 to 3.5 m/s across most of BC, with slightly elevated values of 2.0 to 3.5 m/s in the northern regions. The relationship between mean wind speed and fire occurrence density is less visually apparent than for thermal or moisture variables, suggesting that monthly mean wind speeds may not adequately capture wind dynamics most relevant to fire behaviour. Local wind maxima associated with specific topographic features or specific weather patterns during individual fire events likely exert a stronger influence on fire spread than long-term spatial averages.

2.4 NDVI

The Normalized Difference Vegetation Index (NDVI) was derived from the MODIS Terra Vegetation Indices product (MOD13A2), generated by the Moderate Resolution Imaging Spectroradiometer (MODIS) and distributed by NASA's Land Processes Distributed Active Archive Center (LP DAAC) (Tucker, 1979; Huete et al., 2002). The MOD13A2 product provides 16-day composite NDVI at a spatial resolution of 1 km and has been widely used to characterize vegetation greenness and photosynthetic activity (Didan, 2021). NDVI values are normalized to range from -1 to +1, with higher values indicating denser and healthier vegetation, and were accessed and processed using Google Earth Engine.

In this study, NDVI data covering British Columbia from January 2019 to December 2024 were extracted following the standard provincial extent described in Section 2.1. The original 16-day composite NDVI observations were aggregated to a monthly temporal resolution to align with other covariates and reduce high-frequency noise. Spatial aggregation was performed using zonal statistics, whereby NDVI values were averaged within each 0.25-degree analysis grid. For each grid cell and month, both the mean and standard deviation of NDVI were calculated to capture overall vegetation conditions and within-grid heterogeneity. The original integer NDVI values were rescaled by dividing by 10,000 to restore the standard NDVI range. The resulting dataset consists of 170,424 observations (2,367 grids \times 72 months) and two NDVI-derived variables used for subsequent temporal feature engineering, as summarized in Table 1.

Table 1: NDVI Dataset Schema

Observation	grid_id	Date	Year	Month	NDVI_mean	NDVI_std
1	G_00000	2019-01-01	2019	1		
2	G_00000	2019-02-01	2019	2		
3	G_00000	2019-03-01	2019	3		
:	:	:	:	:	:	:
170,423	G_02366	2024-11-01	2024	11		
170,424	G_02366	2024-12-01	2024	12		

2.5 DEM

Terrain characteristics were derived from the Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM), Version 3, produced by NASA’s Jet Propulsion Laboratory and accessed using Google Earth Engine (NASA JPL, 2013). The SRTM DEM provides global elevation data at a spatial resolution of 30 meters and was generated using radar interferometry during the February 2000 mission (Farr et al., 2007).

Elevation and slope variables capture the vertical relief and terrain steepness that influence fire spread rates and fire intensity, while aspect represents slope orientation that affects solar radiation exposure, fuel moisture conditions, and fire behaviour under prevailing wind regimes (Rothermel, 1972). In this study, three terrain layers were processed for the province of British Columbia:

1. Elevation, representing surface height above sea level in meters;
2. Slope, representing terrain gradient in degrees (0-90°);
3. Aspect, representing slope orientation in degrees (0 - 360°, with 0° indicating north).

Terrain values were extracted at the grid level using the centroid-based point sampling from the 30-meter resolution DEM data, whereby each grid cell was represented by a single elevation, slope, and aspect value of the grid centroids. Additional derived terrain features were constructed, including a categorical slope indicator and sine-cosine transformations of aspect¹. These values provide static environmental predictors that remain temporally constant throughout the study period. In total, six terrain-related variables were extracted per grid cell, resulting in 2,367 observations and illustrated in Table 2.

¹Slope classes were defined informed by the slope–fire spread relationships described by Rothermel (1972): flat($0^\circ - 5^\circ$), gentle($5^\circ - 15^\circ$), moderate($15^\circ - 25^\circ$), steep($> 25^\circ$), and coded as 0, 1, 2, 3 in variable `slope_code` correspondingly.

Table 2: DEM Dataset Schema

Observation	grid_id	elevation	slope	aspect	slope_code	aspect_sin	aspect_cos
1	G_00000						
2	G_00001						
3	G_00002						
:	:	:	:	:	:	:	:
2366	G_02365						
2367	G_02366						

2.6 Landcover

Land cover classification was obtained from the MODIS Land Cover Type product (MCD12Q1, Version 6.1), produced by NASA’s Land Processes Distributed Active Archive Center (LP DAAC) and accessed via Google Earth Engine (M. Friedl & Sulla-Menashe, 2022). Land cover data serve as critical predictors in wildfire risk modelling, as different vegetation types exhibit varying flammability, fuel loads, and susceptibility to fire, and thus play an important role in fire occurrence (Chuvieco et al., 2008).

The MCD12Q1 product provides annual global land cover maps at a spatial resolution of 500 meters, derived from MODIS Terra and Aqua surface reflectance observations using supervised classification algorithms (M. A. Friedl et al., 2010). In this study, the International Geosphere–Biosphere Programme (IGBP) land cover classification scheme (LC-Type1) was adopted (see Table S3 for class definitions).

The 2022 land cover map over British Columbia was selected to represent landscape conditions during the study period, as it corresponds to the midpoint of the analysis window. The land cover data were aggregated to the 0.25-degree analysis grids using a majority resampling approach. For each grid cell, the dominant land cover class was identified based on the most frequently occurring MODIS land cover category within the grid. This procedure yields grid-level static land cover information, resulting in 2,367 observations (one per grid cell) used in subsequent analyses.

3 Feature Engineering

We describe a framework for integrating the multi-source spatiotemporal datasets described in Section 2 into a unified feature set, which serves as the foundation for the statistical analysis and machine learning models employed in this study.

3.1 Fire Labels

Fire incidents are originally recorded at point-level observations, whereas the analysis in this study is conducted on a regular spatial grid. To ensure spatial alignment between the response variables and covariates, fire points were therefore aggregated to grid-level fire occurrence labels. Since multiple fire related timestamps may be available for each incident, including the report, attack, and the put-out date for each fire incident. To reduce potential temporal ambiguity between fire incidents, the earliest date observed was selected to represent the initial fire occurrence.

Fire incidents were then spatially linked to the analysis grids using a point-in-polygon operation, assigning each fire to the corresponding grid cell in which it occurred. The matched fire points were aggregated at the grid-day level, producing a binary indicator that denotes whether at least one fire occurred within a given cell on a specific day. These fire occurrence indicators were merged with the full set of grid-date combinations, and grid-day pairs without recorded fires were explicitly labelled as non-fire events with 0. The resulting datasets comprise 5,188,464 grid-day observations, each associated with a grid identifier, date, and binary wildfire occurrence labels. Fire occurrence was observed in 7,788 grid-day cells over the six-year study period, accounting for only 0.15% of all grid-day observations.

3.2 ERA5 Feature Selection

Among the ERA 5 variables (Table S2), strong positive dependencies were observed describing similar physical processes, particularly within temperature related fields, including mean, maximum and minimum air temperature, dew point temperature, and skin temperature. These variables exhibited pairwise correlation exceeding 0.9, which indicates substantial redundancy. In addition, air temperature showed strong positive correlations with soil temperature at all depth levels, with an average correlation of approximately 0.8. Variables defined across multiple vertical layers also exhibited pronounced internal collinearity, notably among soil temperature layers and volumetric soil water content layers.

Accordingly, a reduced set of ERA5 variables was selected by retaining a single representative variable from each highly correlated group. Specifically, the mean air temperature was preserved to represent the near-surface thermal conditions, while the remaining temperature related variables were excluded. For soil properties, only the top layer soil temperature and volumetric soil water content were retained.

The remaining variables of precipitation, surface runoff, evaporation, wind components, surface pressure, snowfall, and radiation variables were retained from their distinct physical interpretation, and comparatively weaker correlation with other predictors. The final reduced

ERA5 feature set consists of 13 meteorological variables.

3.3 Bad Grids

During feature extraction, a small subset of analysis grids show missing values in one or more environmental covariates. These issues potentially arose from incomplete coverage in the ERA5 reanalysis data and a gap in the DEM data. Specifically, 70 grids contained missing ERA5 variables across the study period, while 8 grids exhibited missing DEM values. Of these, 8 grids overlapped between the data sources, resulting in a total of 73 grids with incomplete covariate information.

As shown in Figure S3, these grids are primarily located along the southwestern coastal region in BC. As discussed in Section 2.2, these regions only experienced sparse isolated fire incidents during the study period. Consequently, the exclusion of these grids is not expected to affect the analysis of wildfire occurrence, and therefore, these grids were removed from the subsequent analyses, resulting in a final dataset comprising 2,294 grid cells.

3.4 Feature Merge

To facilitate consistent spatiotemporal analysis, all datasets were integrated into a unified analysis framework. We first constructed a complete grid-day index by taking the Cartesian product of grid identifiers and the sequence of daily timestamps over the study period. Fire occurrence labels were first merged onto this grid-day panel using matching grid identifiers and calendar dates. Meteorological covariates from the reduced ERA5 dataset were directly aligned to the panel at the daily resolution based on exact grid–date correspondence. Monthly NDVI values were incorporated by replicating each grid’s NDVI value to all days within the corresponding month, matching via grid identifiers and year-month combinations. Static covariates derived from DEM and land cover data were subsequently merged at the grid level and replicated across all dates by grid identifiers, as these attributes remain invariant throughout the study period. The resulting feature set comprises 5,028,448 observations and 29 covariates.

3.5 Features Addition

In addition to the baseline environmental covariates described above, a set of derived and auxiliary features was constructed to capture better persistent environmental conditions and dynamic processes relevant to wildfire occurrence. The additional features are summarized in Table S4, and the final dataset contains 5,028,448 observations and 62 covariates.

Antecedent climate conditions

Lagged temperature values at 1-, 3-, and 7-day intervals were included to represent delayed thermal influences on ignition risk and short-term climate effects. Rolling summaries were further constructed to characterize accumulated conditions, including rolling mean temperature and accumulated precipitation over 7- and 30-day windows. Prolonged drought conditions were explicitly captured by tracking the number of consecutive dry days.

Wind intensity and direction

Wind speed magnitude was computed from the original zonal and meridional wind components, and the maximum wind speed over a 7-day window was included to capture short-term wind events. Wind direction was additionally derived and encoded using sine and cosine transformations to preserve its circular structure.

Seasonal Structure

Seasonal patterns in wildfire occurrence were explicitly encoded using sine and cosine transformations of the calendar month. This allows for smooth cyclical variation across years, and in addition, a binary indicator was included to denote the core wildfire season in British Columbia.

Fire History Effects

Peng et al. (2005) have shown that wildfire events are not independent in time, and that the occurrence of the fire tends to inhibit subsequent fires at the same location, with risk gradually increasing over time following the event. For each grid cell, the number of days was included since the most recent fire and the cumulative historical fire count was computed, and a large sentinel value was assigned to grids that never experienced fire. A dynamic temporal baseline fire risk was further estimated using grid and month specific historical probabilities.

Interaction and extreme event features

Interaction features were introduced to capture combined terrain–climate effects and vegetation conditions relevant to flammability, including elevation–temperature and vegetation–soil moisture interactions. Extreme weather indicators were also incorporated to represent the effects of heatwaves, strong winds, and prolonged drought on wildfire risk.

4 Methodology

The previous section demonstrated that the wildfire occurrence is driven by complex interactions among environmental covariates. Accordingly, we formulate wildfire prediction as a spatiotemporal binary classification problem on a spatial graph, where the objective is to predict next-day fire occurrence for a given grid cell and date.

Our proposed method consists of three main components. First, we construct a spatial graph over the study region of BC using K-nearest neighbours(KNN) to model the geographic proximity among grid cells. Subsequently, each grid cell is represented by a time series vector that represents the local environmental conditions. Finally, these inputs are processed by a Spatiotemporal Graph Neural Network (ST-GNN) that integrates two-layer graph convolutions for spatial aggregation, simultaneously with a two-layer bidirectional Long Short-Term Memory (LSTM) (Hochreiter & Schmidhuber, 1997) for temporal modelling. The entire architecture is trained using Focal Loss to address extreme class imbalance in the data, and aims to predict next day wildfire occurrence probabilities for all grid cells.

4.1 Datasets Partition

The dataset spans the years 2019 to 2024 for the study region of British Columbia. To avoid data leakage and to simulate the real-world forecasting scenarios, we partition the dataset chronologically into three subsets. All subsets share the same spatial graph structure and feature representation. The training set covers the period January 1, 2019, to December 31, 2022, providing 4 complete years of data for model fitting. The validation set spans from January 1, 2023, to December 31, 2023, and is used for hyperparameter tuning, early stopping, and learning rate scheduling. The test set comprises the period from January 1, 2024, to December 31, 2024, and enabling evaluation of model performance on unseen conditions. In terms of sample size, the partitioning results in 3,351,534 daily grid observations for training, 837,310 for validation, and 839,604 for testing.

4.2 Spatial Graph

To further refine the modelling framework, we aim to capture the geographic proximity and potential interaction relevant to wildfire spread using an undirected spatial graph. The spatial graph structure enables the nodes (grid cells) to aggregate information from neighbouring regions where fire may spread by capturing the proximity relationship between grid cells.

We first construct a weighted spatial graph $G = (V, E)$ over the study region of British Columbia. Each grid cell is treated as a node $v_i \in V$, and Edges are established based on

geographic distance between cell centroids. We adopt a K-nearest neighbours strategy with $K = 8$, as 8 neighbours approximate the immediate surrounding cells in a regular grid. For each grid cell v_i with centroid coordinates $(\text{lon}_i, \text{lat}_i)$, we define the set of neighbours as,

$$\mathbf{N}(v_i) = \{v_j : d(v_i, v_j) \leq d_k(v_i), j \neq i\}, \quad (1)$$

where $d(v_i, v_j)$ is the Euclidean distance between grid centroids, and $d_k(v_i)$ is the distance to the k-th nearest neighbour of v_i . Additionally, we impose a maximum distance threshold of $d_{\max} = 35$ km to prevent linking cells that are geographically unlikely to interact in terms of fire spread. Thus, the edge set is defined as,

$$\mathbf{E} = \{(v_i, v_j) : v_j \in \mathbf{N}(v_i) \text{ and } d(v_i, v_j) \leq d_{\max}\} \quad (2)$$

Next, to reflect the decay of fire spread likelihood with distance, and each edge $(v_i, v_j) \in E$ is assigned a weight that decays with distance. We use a Gaussian Kernel to define (Defferrard et al., 2016),

$$w_{ij} = \exp\left(-\frac{d(v_i, v_j)^2}{2\sigma^2}\right), \quad (3)$$

where σ controls the decay rate of spatial influence. In this study, we use $\sigma = 15$ km based on the typical fire spread distance observed in the study region. The resulting weighted adjacency matrix $W \in \mathbb{R}^{N \times N}$ is normalized to ensure numerical stability using symmetric normalization (Kipf & Welling, 2017),

$$\tilde{A} = D^{-\frac{1}{2}} W D^{-\frac{1}{2}}, \quad (4)$$

where D is the degree matrix with $D_{ii} = \sum_j W_{ij}$.

4.3 ST-GNN Architecture

Building upon the spatial graph defined in Section 4.2, we propose a Spatial Temporal Graph Neural Network (ST-GNN) (Yu et al., 2018), which is designed to jointly model the spatial dependencies and temporal dynamics inherent to wildfire occurrence. The model processes the historical sequence of graph-structured data to predict fire probabilities for the next day. The architecture mainly comprises four core components, an input embedding layer, a spatial graph convolutional layer, a temporal sequence modelling layer, and a prediction output layer.

4.3.1 Input Embedding Layer

The input embedding layer transforms the raw feature vector $x_i^{(t)}$ for each node v_i at time t by projecting it into a higher-dimensional latent space (we use $h = 64$ in this study) to enhance the model's representational capacity. This is achieved through a shared fully-connected layer followed by a Rectified Linear Unit (ReLU) activation (Nair & Hinton, 2010),

$$H^{(t)} = \text{ReLU}(X^{(t)}W_{emb} + b_{emb}), \quad (5)$$

where $W_{emb} \in \mathbb{R}^{d \times h}$ is the embedding weight matrix, $b_{emb} \in \mathbb{R}^h$ is the bias vector, h is the hidden dimension. This operation is applied independently to each time step $t \in \{t - \tau + 1, \dots, t\}$ in the input window, where τ is the manually selected length of the input time window. This further transform the input sequence $X^{(t-\tau+1:t)}$ into a sequence of embedded node representations $H^{(t-\tau+1)}, \dots, H^{(t)}$.

4.3.2 Spatial Graph Convolution Layer

To capture spatial dependencies among neighbouring grid cells, we employ two stacked graph convolutional network (GCN) layers. Each layer aggregates features from a node's neighbours according to the normalized adjacency matrix \tilde{A} from Section 4.2. Following Kipf & Welling (2017), the graph convolution operation for a single time step is defined as,

$$H^{(t,l+1)} = \text{ReLU}(\tilde{A}H^{(t,l)}W^{(l)}) \quad (6)$$

where $H^{(t,l)} \in \mathbb{R}^{N \times h}$ is the node representation at layer l and time t , and $W^{(l)} \in \mathbb{R}^{h \times h}$ is the learnable weight matrix for layer l . We stack $L = 2$ graph convolutional layers to capture multi-hop spatial dependencies relevant to fire spread, which allows the model to integrate information from both immediate and second-order neighbours. After each GCN layer, we apply batch normalization (Ioffe & Szegedy, 2015) to (6) to stabilize training,

$$H^{(t,l+1)} = \text{BatchNorm}\left(\text{ReLU}\left(\tilde{A}H^{(t,l)}W^{(l)}\right)\right) \quad (7)$$

The graph convolution is applied independently to each time step t in the input sequence, the output is a set of spatially aggregated representations $\{Z^{(t-\tau+1)}, \dots, Z^{(t)}\}$ where $Z^{(t)} = H^{(t,L)}$ denotes the output of the final GCN layer at time t .

4.3.3 Temporal Sequence Modelling Layer

To capture temporal dependencies and accumulation effects, we employ a two-layer bidirectional Long Short-Term Memory (LSTM) network (Hochreiter & Schmidhuber, 1997). This allows the temporal representation to incorporate information from the entire input window in both forward and backward directions, and selectively retain or forget information from different timesteps. For each grid cell v_i , we extract its temporal feature sequence z_i ,

$$z_i = \left[Z_i^{(t-\tau+1)}, Z_i^{(t-\tau+2)}, \dots, Z_i^{(t)} \in \mathbb{R}^{\tau \times h} \right] \quad (8)$$

The bidirectional LSTM processes this sequence in both forward and backward directions, allowing the temporal representation to incorporate information from the entire historical window. We concatenate the final hidden states from the forward and backward passes, where $h_i^{\text{bi}} = [h_{\tau}^{\text{forward}}; h_1^{\text{backward}}] \in \mathbb{R}^{2h}$. We then apply a linear projection to obtain a temporal representation $\tilde{h}_i \in \mathbb{R}^h$, which encodes information from all τ historical time steps with both forward and backward context. This operation is performed in parallel for all N grid cells, yielding the final spatiotemporal representation $H_{LSTM} \in \mathbb{R}^{N \times 2h}$. Next, we apply dropout regularization (Srivastava et al., 2014) to improve generalization and prevent overfitting. We compute $H_{LSTM} = \text{Dropout}(H_{LSTM}, p = 0.3)$, where dropout randomly sets 30% of the activations to zero during training.

4.3.4 Prediction Output Layer

The final spatiotemporal representations for all nodes, $H_{LSTM} \in \mathbb{R}^{N \times 2h}$, are passed through a two-layer fully-connected network to produce logits for each grid cell,

$$\begin{aligned} H_{pred} &= \text{ReLU}(H_{LSTM}W_1 + b) \\ H_{pred} &= \text{Dropout}(H_{pred}, p = 0.3) \\ \hat{y} &= H_{pred}W_2 + b_2 \end{aligned}$$

where $W_1 \in \mathbb{R}^{2h \times \frac{h}{2}}$, $W_2 \in \mathbb{R}^{\frac{h}{2} \times 1}$, $b_1 \in \mathbb{R}^{\frac{h}{2}}$, and $b_2 \in \mathbb{R}$ are learnable parameters. Finally, the logit \hat{y} for cell v_i is converted to a wildfire occurrence probability by applying the sigmoid function,

$$p_i = \sigma(\hat{y}_i) = \frac{1}{1 + e^{-\hat{y}_i}} \quad (9)$$

where $p_i \in (0, 1)$ represents the predicted probability of the occurrence of wildfires in the grid cell v_i on day $t + 1$.

4.4 Loss Function

Given the extreme class imbalance present in the wildfire dataset, with approximately 0.15% fire observations and 99.85% non-fire observations, we adopt the Focal Loss Lin et al. (2017) as the training objective. Focal Loss dynamically scales the contribution of each sample based on the model’s prediction confidence, it effectively down-weighting well-classified examples while focusing learning on hard and misclassified cases. The loss for a single prediction is defined as,

$$\mathcal{L}_{FL}(p, y) = -\alpha_t(1 - p_t)^\gamma \log(p_t) \quad (10)$$

where

$$p_t = \begin{cases} p & \text{if } y = 1 \\ 1 - p & \text{if } y = 0 \end{cases}, \quad \alpha_t = \begin{cases} \alpha & \text{if } y = 1 \\ 1 - \alpha & \text{if } y = 0 \end{cases}$$

The hyperparameter $\gamma \geq 0$ reduces the relative loss for well-classified examples, where Focal Loss is equivalent to standard cross-entropy when $\gamma = 0$. As γ increases, the effect of the modulating factor $(1 - p_t)^\gamma$ increases. In this study, we use $\gamma = 2$ following the empirical recommendations in Lin et al. (2017).

The other hyperparameter $\alpha \in [0, 1]$ addresses class imbalance by assigning a higher weight to the minority class. We set $\alpha = 0.9$ to reflect the extreme imbalance in our dataset to balance the contribution of fire and non-fire classes to the total loss and to increase the relative importance of the positive fire class.

5 Results

We conducted a systematic study across five model configurations to evaluate the sensitivity of the proposed wildfire ST-GNN to temporal context, training duration, and feature engineering. The standard configuration employed a 14-day historical input window and was trained for 20 epochs using the full set of 63 engineered features. To assess the impact of temporal context length, we modified the input sequence windows to 21 days and 7 days while keeping all other standard settings unchanged. To evaluate the effect of training duration, a third variant increased the number of training epochs from 20 to 50 while training

the 14-day input window and full feature set of the standard model. Finally, we trained a reduced model using only raw environmental covariates, excluding all engineered interaction features described in Section 3.5. We want to assess whether feature construction provides effective predictive value beyond the raw observational inputs.

5.1 Threshold Optimization

While binary classification models conventionally adopt a probability threshold of 0.5, this default choice is inappropriate for wildfire occurrence prediction due to the extreme class imbalance and the rarity of wildfire events. To address the issue, we performed a threshold optimization analysis using the standard model configuration.

Specifically, we evaluate thresholds ranging from 0.05 to 0.5 with a step size of 0.05. At each threshold, classification performance was assessed on the 2024 test set, and the optimal threshold was selected by maximizing the F1 score, which balances precision and recall metrics.

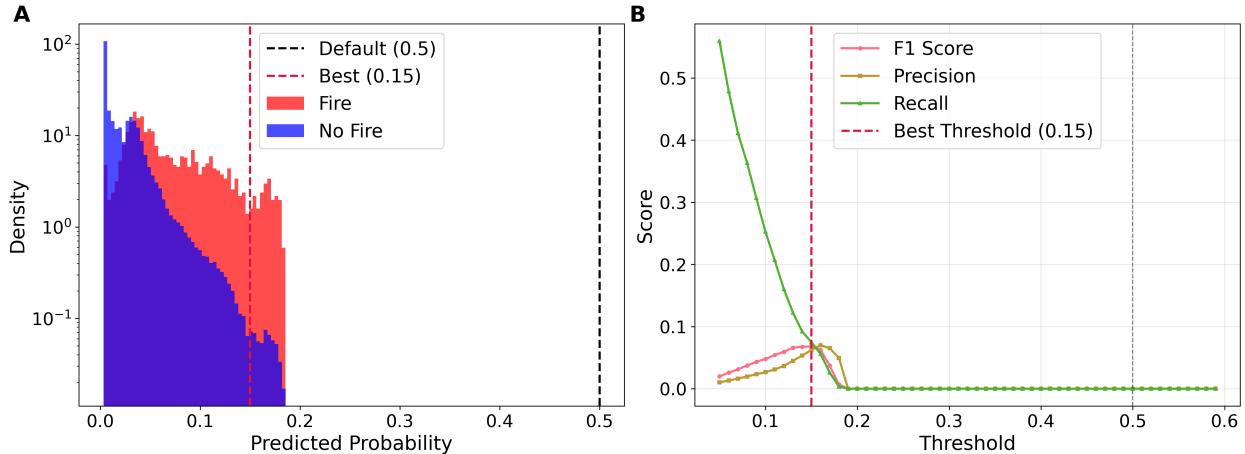


Figure 4: (A) Distribution of predicted wildfire probabilities for fire and non-fire cases on a logarithmic density scale; (B) Trade-off between precision, recall, and F1-score across classification thresholds.

Figure 4A illustrates the distribution of predicted probabilities for fire and non-fire instances on the 2024 test set. The majority of predictions, including many true fire events in the red distribution, are concentrated in the low probability range from 0 to 0.2. In contrast, the non-fire instances in the blue distribution show a stronger concentration near zero. Notably, all predicted probabilities fall below the conventional threshold of 0.5, resulting in the model classifying every instance as no fire regardless of the underlying risk level.

As shown in Figure 4B, at a very low threshold below 0.1, the model can correctly identify most fire events as recall approaches its maximum value of 0.56. However, it comes at the cost of extremely low precision of 0.01 due to massive false positive rates. As the threshold increases beyond 0.2, Recall declined rapidly while precision improved only marginally. The best trade-off is achieved at a threshold of 0.15 with the maximum F1 score of 0.0679, which we adopt in all subsequent analyses.

5.2 Basic Classification

We conducted predictions on the 2024 test set across all five model configurations, and the corresponding classification results are shown in Table 3. All models evaluate at the optimal threshold of 0.15 identified in Section 5.1, with a total of 839,604 test observations including 838,212 non-fire events (negative) and 1,392 fire events (positive).

Across all configurations, the models achieved seemingly high accuracy values exceeding 98%. However, accuracy is not an informative performance metric under such extreme class imbalance. Given that fire events account for only 0.17% of the test set, a naive classifier that predicts no fire for every instance would already achieve 99.83% accuracy. We focus on the trade-off between evaluation metrics and confusion matrix statistics, which better reflect model behaviour in our wildfire prediction.

Table 3: Classification performance and confusion matrix (Figure S5) statistics by model configuration.

Model	Accuracy	Precision	Recall	TN	FP	FN	TP	False Alarm Ratio (FP/TP)
Standard	0.9824	0.0265	0.2525	777,134	12,906	1,039	351	36.8
Seq21	0.9930	0.0362	0.1165	785,728	4,312	1,228	162	26.6
Seq7	0.9821	0.0278	0.2698	776,914	13,126	1,015	375	35.0
Reduced	0.9867	0.0299	0.2094	780,586	9,454	1,099	291	32.5

Note: The Epo50 model yields identical confusion matrix statistics to the Standard model and is omitted.

Among the temporal variants, the 21-day sequence model (Seq21) adopts the most conservative prediction strategy. It produces only 4,312 false alarms while achieving the highest precision of 0.0362 among all configurations. However, this approach comes at the expense of sensitivity, yielding the lowest recall of 0.1165 and detecting only 162 fire events. In contrast, the 7-day sequence model (Seq) shows the most aggressive prediction behaviour. By responding more reactively to the recent fire weather conditions, it generates the largest number of false alarms of 13,126, but also detects the highest number of 375 fire events. This

achieved the highest recall of 0.2698 among all models. These results suggest that shorter temporal context contains more direct information relevant to the wildfire events as reflected in the higher sensitivity, while longer historical windows impose stricter requirements for fire prediction and result in fewer but more selective alerts.

The reduced model that relies solely on the raw environment covariates sacrifices its sensitivity relative to the standard configuration. It achieved a recall of 0.2094, but it produces 3,452 fewer false positives and also detects 60 fewer true events. This indicates that the fire risk detection is reduced when engineered interaction features are excluded.

The standard configuration model with a 14-day temporal window and the full feature set represents the most balanced model between sensitivity and false alarm control. It detected 351 true fire events that achieved a second highest recall of 0.2525, while incurring the highest false alarm ratio of 36.8%. This implies that the standard model generates 36.8 false alarms for every true fire detected on average, and potentially reducing the likelihood of wildfire occurrence from early investigation or intervention is performed.

5.3 Temporal Analysis

We conducted a temporal performance analysis to assess whether model behaviour varies across the annual wildfire cycle, with particular emphasis on differences between the fire season from April to September and the non-fire season from October to March. The results indicate that annual performance metrics masked the strong seasonal heterogeneity of models, while the architectural differences have consistent effects on performance across time.

Specifically, performance metrics increase sharply during summer months, with all models exhibiting their strongest performance between July and August. Figure 5 shows the Seq7 model achieves the highest monthly recall peaked at 0.47 in July, effectively identifying nearly half of all fire events during the most active period. This strong performance persists into August and attains the highest recall of 0.30. However, this seasonal strength is partially masked by the annual average recall of 0.27 from the inclusion of low activity months. The standard configuration demonstrated a more balanced temporal response by retaining moderate sensitivity beyond the peak fire months. Although its recall is slightly lower than Seq7 during July and August, it achieved the highest recall in September of 0.05 while all other models dropped to near-zero during this period.

In contrast, all performance metrics collapse to exactly zero throughout the non-fire season across all model configurations. As summarized by the confusion matrices in Table S6, no true fire events were detected despite the presence of a small number of false positives

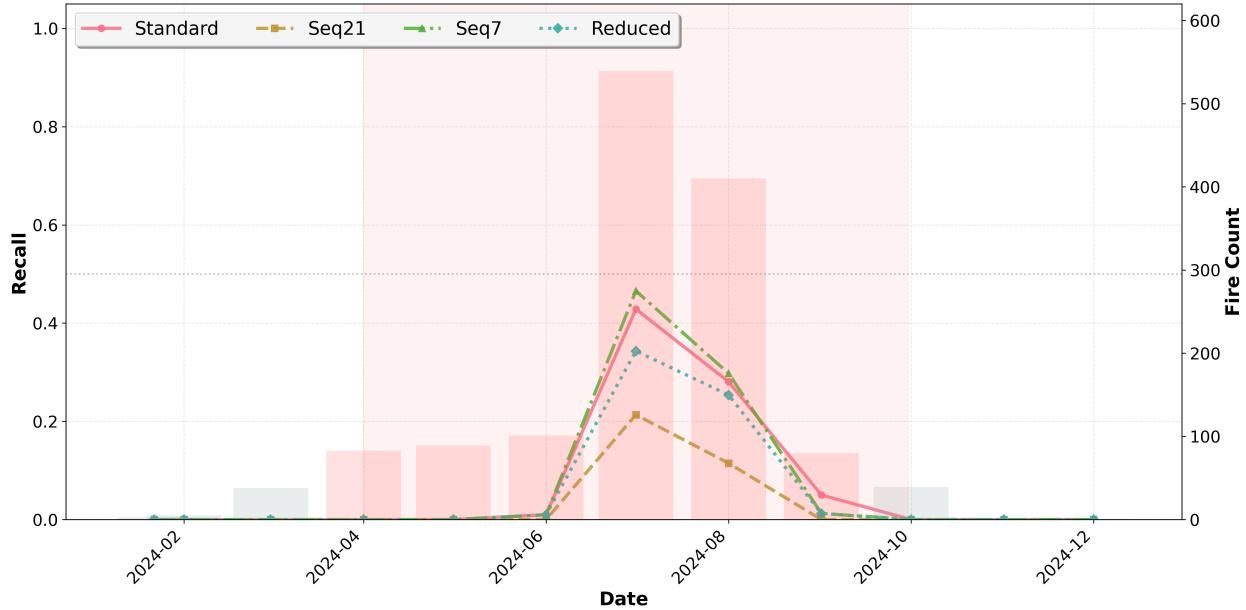


Figure 5: Temporal recall of wildfire prediction models aggregated by month

across models. This behaviour is potentially attributable to the extremely rare fire occurrence that arises from localized or human-caused conditions that are weakly correlated with the environmental drivers used by the models.

These findings suggest that a seasonal adaptive thresholding strategy may be more appropriate than an annual fixed classification threshold. During the peak fire season months, a lower threshold such as 0.1 may be preferable to allow higher false alarm rates in exchange for improved early detection. Conversely, maintaining the standard threshold of 0.15 during shoulder months such as May and October. This is appropriate as fire activity is lower, and excessive false alarms may impose unnecessary operational burden during early investigation and response.

5.4 Spatial Analysis

We next conducted a spatial analysis of wildfire risk across British Columbia during the 2024 fire season. We want to evaluate whether different model configurations reproduce consistent spatial patterns with distinct regional risk structures as the original data.

As shown in Figure 6, all model configurations successfully captured the dominant spatial structure of wildfire risk in British Columbia, consistently assigning higher predicted probabilities to the interior and southeastern regions than the coastal and northern areas. Rather than identifying isolated high risk grid cells, all models produce continuous spatial

risk gradients, with predicted probabilities varying smoothly across longitude.

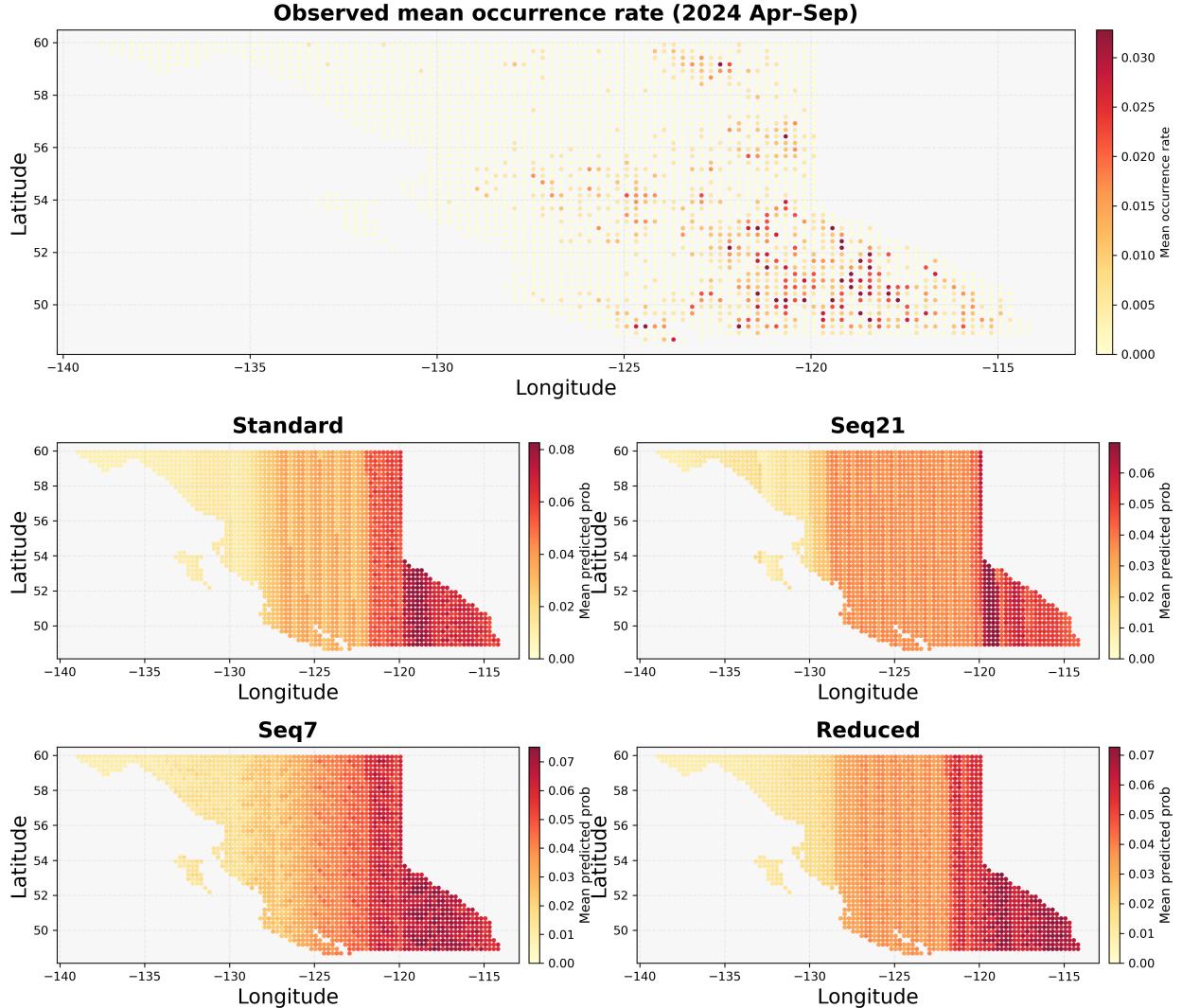


Figure 6: Spatial distribution of wildfire occurrence and model-predicted risk during the 2024 fire season from April to September

The standard model produces a spatial risk surface that most closely aligns with the observed distribution. It captures moderate spatial gradients while preserving meaningful spatial variation without over-amplifying regional risk. High risk areas remain spatially coherent but not overly concentrated, and the model clearly differentiates the lower risk northern regions from higher risk southern and southeastern regions. The Seq21 model generates a reduced local variability surface with more conservative and homogeneous moderate risk. The higher risk regions are concentrated within a narrower band of southeastern grids.

In contrast, the Seq7 model yields the most spatially aggressive risk surface. Its predic-

tions are characterized by sharper gradients and stronger contrast between adjacent regions, where high predicted probabilities extend more broadly across the southern interior. Finally, the reduced model shows intermediate behaviour between the Standard and Seq7 configuration. It assigns a more moderate risk across interior regions while producing a relatively high risk classification in the southeastern areas.

6 Discussion

In this study, we demonstrated a framework to predict the wildfire occurrence in British Columbia by integrating historical wildfire point records with the spatial context within a grid-based analysis structure. We pre-processed the multi-source environmental datasets, including ERA5 weather variables, terrain characteristics, vegetation indices, and land cover information, and subsequently constructed a unified spatiotemporal feature set for graph-based modelling. This was followed by a feature engineering strategy that systematically incorporates both persistent environmental conditions and dynamic processes relevant to wildfire occurrence. We explicitly constructed geographical proximity between grid cells using KNN, while modelled temporal dependencies through a bidirectional LSTM architecture across four model configurations. The results demonstrate that the proposed models successfully reproduce the spatial distribution of wildfire risk in British Columbia. They exhibit strong predictive performance during the peak fire season, while performance decays during periods with negligible fire activity. These findings further highlight the influence of input sequence length and environmental covariates on model performance, and illustrate the effectiveness of explicit modelling spatiotemporal dependencies in wildfire occurrence prediction.

We encountered significant computational constraints that influenced our experimental design of this study. Training the proposed ST-GNN architecture on 2,367 spatial nodes and 5 million temporal observations using a 2021 M1 iMac with 8GB unified memory required nearly 2 hours per epoch under the standard model configurations. As a result, completing the full 20 epochs training cycle required approximately 40 hours, which optimizing hyper parameters through automated search methods becomes infeasible. Instead, we adopted a targeted experimental strategy that focused on a limited set of selected model configurations, which allowed for meaningful comparisons of architectural and performance.

Despite the promising results, several limitations of the current study should be acknowledged. First, our analysis was conducted at a spatial resolution of 0.25 degrees, that approximately 25 km x 25 km, which inevitably obscures localized heterogeneity in fire ignition and wildfire spread. While this resolution is computationally efficient and consistent

with the native scale of ERA5 reanalysis data, it fails to capture fine-scale topographic variation, microclimatic effects, and localized human activities that strongly influence the wildfire process. Second, land cover was represented using a single static map from 2022, which does not account for interannual landscape changes arising from environmental changes or prior wildfire disturbance. Although the inclusion of fire history features, such as days since last fire, these variables do not fully capture the ecological recovery post fire or temporal change in fuel structure and flammability.

Moreover, our model was trained on data from 2019 to 2022, with 2023 serving as the validation set and 2024 as the test set. However, as discussed in Section 2.2, this temporal partitioning introduces a temporal distribution shift. The baseline years of 2019 and 2020 recorded relatively modest wildfire activity, while the extreme fire regime began in 2021 and persisted in subsequent years. Consequently, the model may have been insufficiently exposed to the extreme fire weather conditions and spatiotemporal patterns that dominate in validation and test periods during training. This may potentially contribute to the conservative prediction behaviour observed in the results, and increase the false negative rate during peak fire months.

Furthermore, a notable limitation of this study is the absence of direct performance comparisons with traditional machine learning baselines such as XGBoost, Random Forest, or Gradient Boosting Machines, which have been widely adopted in wildfire prediction literature due to time constraints. Such comparisons are particularly important because these baseline methods typically treat each grid cell as an independent prediction instance and do not explicitly model spatial dependencies. Inclusion of these benchmarks would help to quantify the performance gains attributable to the proposed graph-based framework and provide stronger empirical evidence for the explicit spatial structure in rare event prediction.

We identify three primary directions for future research. First, the current binary wildfire occurrence framework could be extended to a multi-task learning setting that jointly predicts fire occurrence, burned area and fire duration. This would provide more comprehensive risk assessments and may improve overall prediction accuracy and interpretability. Moreover, we intend to expand the temporal training window from the current six-year period to two decades or longer. This would allow the model to capture a broader range of climate variability and fire weather regimes, in particular, incorporating extreme fire seasons such as 2017 and 2018 in training. Finally, the proposed modelling framework could be transferred and adopted to other Canadian provinces, including Alberta, Ontario, and Saskatchewan. Thereby reduces the cost of developing region-specific models and contributes toward a unified Canadian wildfire prediction system.

References

- Abatzoglou, J. T., & Williams, A. P. (2016). Impact of anthropogenic climate change on wildfire across western us forests. *Proceedings of the National Academy of sciences*, 113(42), 11770–11775.
- Anselin, L. (1995). Local indicators of spatial association—lisa. *Geographical Analysis*, 27(2), 93–115.
- BC Wildfire Service. (2025). *Wildfire season summary*. <https://www2.gov.bc.ca/gov/content/safety/wildfire-status/about-bcws/wildfire-history/wildfire-season-summary>. (Accessed: 2025-12-01)
- Canadian Forest Service. (2024). *National burned area composite (nbac)*. Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre. Retrieved from <https://cwfis.cfs.nrcan.gc.ca> (Accessed 2025-11-15)
- Canadian Interagency Forest Fire Centre. (2023). *National fire situation report 2023*. Retrieved from <https://cifffc.net/situation/2023-10-06> (Accessed: 2025-11-28)
- Canadian Interagency Forest Fire Centre. (2025). *National fire situation report 2025*. Retrieved from <https://cifffc.net/situation/2025-09-30> (Accessed: 2025-11-28)
- Church, M., & Ryder, J. M. (2010). Physiography of british columbia. *Compendium of forest hydrology and geomorphology in British Columbia*, 1, 17–45.
- Chuvieco, E., Giglio, L., & Justice, C. (2008). Global characterization of fire activity: toward defining fire regimes from earth observation data. *Global change biology*, 14(7), 1488–1502.
- Defferrard, M., Bresson, X., & Vanderghenst, P. (2016). Convolutional neural networks on graphs with fast localized spectral filtering. *Advances in neural information processing systems*, 29.
- Didan, K. (2021). *MODIS/Terra Vegetation Indices 16-Day L3 Global 1km SIN Grid V061*. NASA EOSDIS Land Processes Distributed Active Archive Center. Retrieved from <https://doi.org/10.5067/MODIS/MOD13A2.061>
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., ... others (2007). The shuttle radar topography mission. *Reviews of geophysics*, 45(2).

- Friedl, M., & Sulla-Menashe, D. (2022). Modis/terra+ aqua land cover type yearly l3 global 500 m sin grid v061, nasa eosdis land processes distributed active archive center [data set]. DOI: <https://doi.org/10.5067/MODIS/MCD12Q1, 61>.
- Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). Modis collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote sensing of Environment*, 114(1), 168–182.
- Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24(3), 189–206.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27. doi: 10.1016/j.rse.2017.06.031
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... others (2020). The era5 global reanalysis. *Quarterly journal of the royal meteorological society*, 146(730), 1999–2049.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735–1780.
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the modis vegetation indices. *Remote sensing of environment*, 83(1-2), 195–213.
- Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning* (pp. 448–456).
- Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. Retrieved from <https://arxiv.org/abs/1609.02907>
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the ieee international conference on computer vision* (pp. 2980–2988).
- Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37(1–2), 17–23.

- Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (icml-10)* (pp. 807–814).
- NASA JPL. (2013). *NASA Shuttle Radar Topography Mission Global 1 Arc Second*. NASA EOSDIS Land Processes Distributed Active Archive Center. Retrieved from <https://doi.org/10.5067/MEaSUREs/SRTM/SRTMGL1.003> doi: 10.5067/MEaSUREs/SRTM/SRTMGL1.003
- Peng, R. D., Schoenberg, F. P., & Woods, J. A. (2005). A space–time conditional intensity model for evaluating a wildfire hazard index. *Journal of the American Statistical Association*, 100(469), 26–35.
- Rey, S. J., Anselin, L., et al. (2019). esda: Exploratory spatial data analysis in python. *Journal of Open Source Software*, 4(42), 1227. doi: 10.21105/joss.01227
- Rothermel, R. C. (1972). *A mathematical model for predicting fire spread in wildland fuels* (Vol. 115). Intermountain Forest & Range Experiment Station, Forest Service, US
- Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., & Monfardini, G. (2008). The graph neural network model. *IEEE transactions on neural networks*, 20(1), 61–80.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1), 1929–1958.
- Statistics Canada. (2021). *2021 census boundary files: Provinces and territories*. <https://www12.statcan.gc.ca/census-recensement/2021/geo/sip-pis/boundary-limites/index2021-eng.cfm>. (Accessed: 2025-11-20)
- Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote sensing of Environment*, 8(2), 127–150.
- Wagner, C. v. (1987). *Development and structure of the canadian forest fire weather index system*. (No. 35).
- Wotton, B. M. (2009). Interpreting and using outputs from the canadian forest fire danger rating system in research applications. *Environmental and ecological statistics*, 16(2), 107–131.
- Yu, B., Yin, H., & Zhu, Z. (2018). Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. *arXiv preprint arXiv:1709.04875*.

Supplementary Material

6.1 Software availability

The code for the analysis are available at https://github.com/Henzhwang/Wildfire_Prediction_STGNN.git.

6.2 Tables and Figures

Table S1: Description of wildfire point dataset variables obtained from the Canadian Wildland Fire Information System (CWFIS) DataMart

Variable	Description	Type
NFDBFIREID	National Fire Database unique identifier	Identifier
FIRE_ID	Provincial wildfire identifier	Identifier
FIRENAME	Assigned wildfire name (if available)	String
SRC_AGENCY	Reporting or managing agency	Categorical
NAT_PARK	Indicator of national park jurisdiction	Binary
LATITUDE	Latitude of wildfire ignition point	Numeric
LONGITUDE	Longitude of wildfire ignition point	Numeric
YEAR	Calendar year of wildfire occurrence	Integer
MONTH	Month of wildfire occurrence	Integer
DAY	Day of wildfire occurrence	Integer
REP_DATE	Date wildfire was officially reported	Date
ATTK_DATE	Date of initial fire attack	Date
OUT_DATE	Date wildfire was declared out	Date
ACQ_DATE	Date wildfire record was acquired	Date
SIZE_HA	Final wildfire size in hectares	Numeric
CAUSE	Primary wildfire cause classification	Categorical
CAUSE2	Secondary or detailed cause classification	Categorical
FIRE_TYPE	Wildfire type classification	Categorical
RESPONSE	Fire response classification	Categorical
PROTZONE	Fire protection zone designation	Categorical
PREScribed	Indicator of prescribed or planned burn	Binary
MORE_INFO	Additional descriptive information	String
CFS_NOTE1	Canadian Forest Service annotation	String
CFS_NOTE2	Canadian Forest Service annotation	String
Large_Fire	Indicator of large wildfire event	Binary
layer	GIS layer identifier	Metadata
omit	Internal data filtering flag	Metadata
geometry	Point geometry representing wildfire location	Spatial

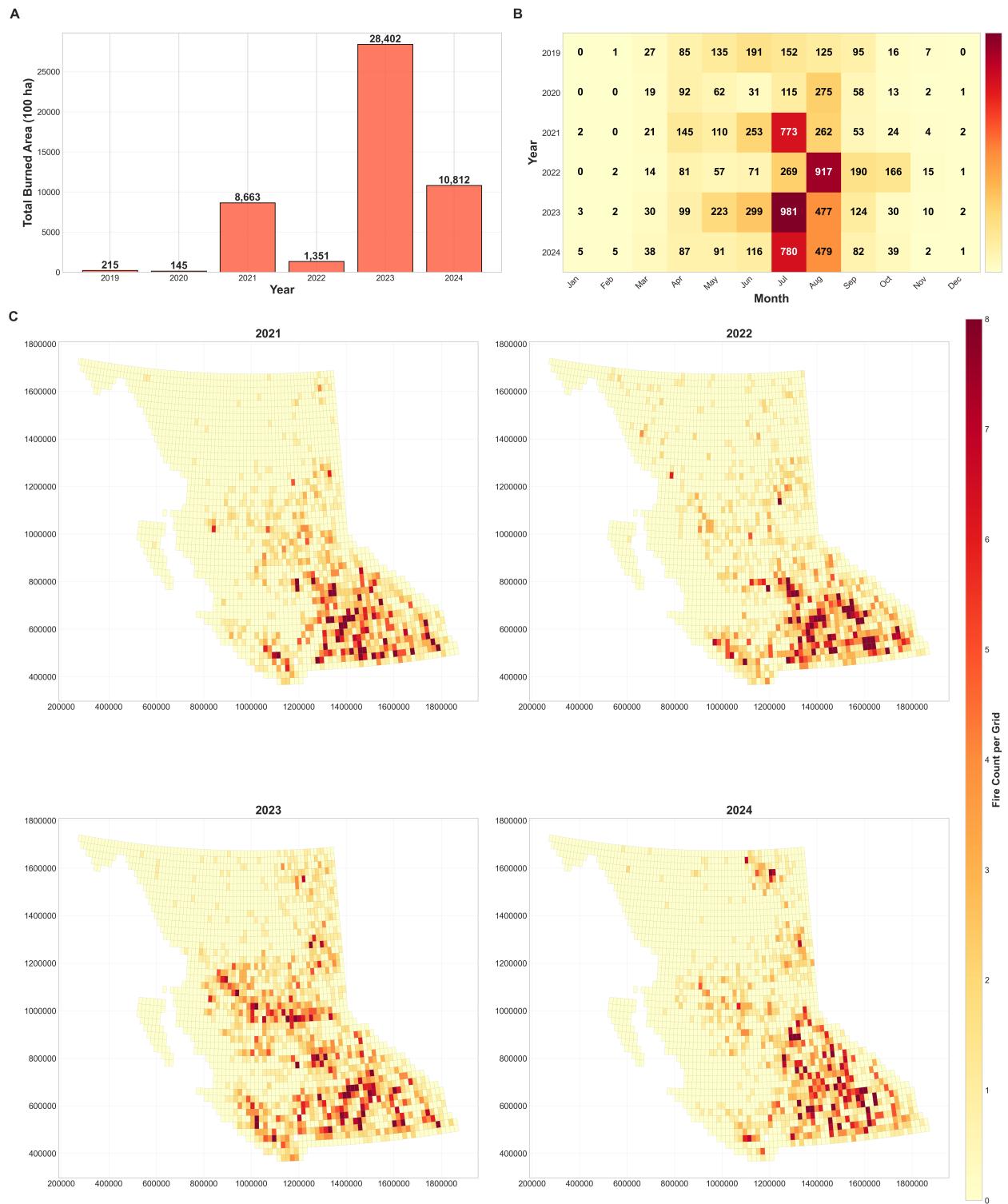


Figure S1: (A) Annual total burned area (km^2) in British Columbia from 2019–2024; (B) Year–month heatmap of wildfire counts illustrating seasonal variation; (C) Spatial distribution of grid-level wildfire counts for 2021–2024, shown using a common colour scale to enable annual comparison. All maps are displayed in the Albers Equal Area projection; axes represent easting and northing in meters.

Table S2: Description of ERA5 weather variables accessed via Google Earth Engine

Variable	Description	Category
temperature_2m	Mean air temperature at 2 meters above ground (K)	Temperature
temperature_2m_max	Daily maximum air temperature at 2 meters above ground (K)	Temperature
temperature_2m_min	Daily minimum air temperature at 2 meters above ground (K)	Temperature
dewpoint_temperature_2m	Dew point temperature at 2 meters above ground, indicating atmospheric moisture (K)	Temperature
skin_temperature	Land surface (skin) temperature (K)	Temperature
total_precipitation_sum	Total accumulated precipitation, including rain and snow (m)	Precipitation
surface_runoff_sum	Total surface runoff generated by precipitation (m)	Hydrology
u_component_of_wind_10m	Zonal (east–west) wind component at 10 meters above ground (m/s)	Wind
v_component_of_wind_10m	Meridional (north–south) wind component at 10 meters above ground (m/s)	Wind
surface_pressure	Atmospheric pressure at the surface (Pa)	Pressure
total_evaporation_sum	Total accumulated evaporation from the surface (m)	Hydrology
surface_solar_radiation_downwards_sum	Total downward solar (shortwave) radiation received at the surface (J/m^2)	Radiation
surface_thermal_radiation_downwards_sum	Total downward thermal (longwave) radiation received at the surface (J/m^2)	Radiation
soil_temperature_level_1	Soil temperature at 0–7 cm depth (K)	Soil Temperature
soil_temperature_level_2	Soil temperature at 7–28 cm depth (K)	Soil Temperature
soil_temperature_level_3	Soil temperature at 28–100 cm depth (K)	Soil Temperature
soil_temperature_level_4	Soil temperature at 100–289 cm depth (K)	Soil Temperature

Continued on next page

Variable	Description	Category
volumetric_soil_water_layer_1	Volumetric soil water content at 0–7 cm depth (m ³ /m ³)	Soil Moisture
volumetric_soil_water_layer_2	Volumetric soil water content at 7–28 cm depth (m ³ /m ³)	Soil Moisture
volumetric_soil_water_layer_3	Volumetric soil water content at 28–100 cm depth (m ³ /m ³)	Soil Moisture
volumetric_soil_water_layer_4	Volumetric soil water content at 100–289 cm depth (m ³ /m ³)	Soil Moisture
snowfall_sum	Total accumulated snowfall expressed as water equivalent (m)	Snow
snow_depth_water_equivalent	Water-equivalent depth of accumulated snowpack (m)	Snow

Table S3: IGBP Land Cover Classes (MODIS MCD12Q1 LC_Type1)

Class Number	Category
0	Water
1	Evergreen needleleaf forest
2	Evergreen broadleaf forest
3	Deciduous needleleaf forest
4	Deciduous broadleaf forest
5	Mixed forests
6	Closed shrublands
7	Open shrublands
8	Woody savannas
9	Savannas
10	Grasslands
11	Permanent wetlands
12	Croplands
13	Urban and built-up
14	Cropland/natural vegetation mosaic
15	Snow and ice
16	Barren or sparsely vegetated
17	Water bodies

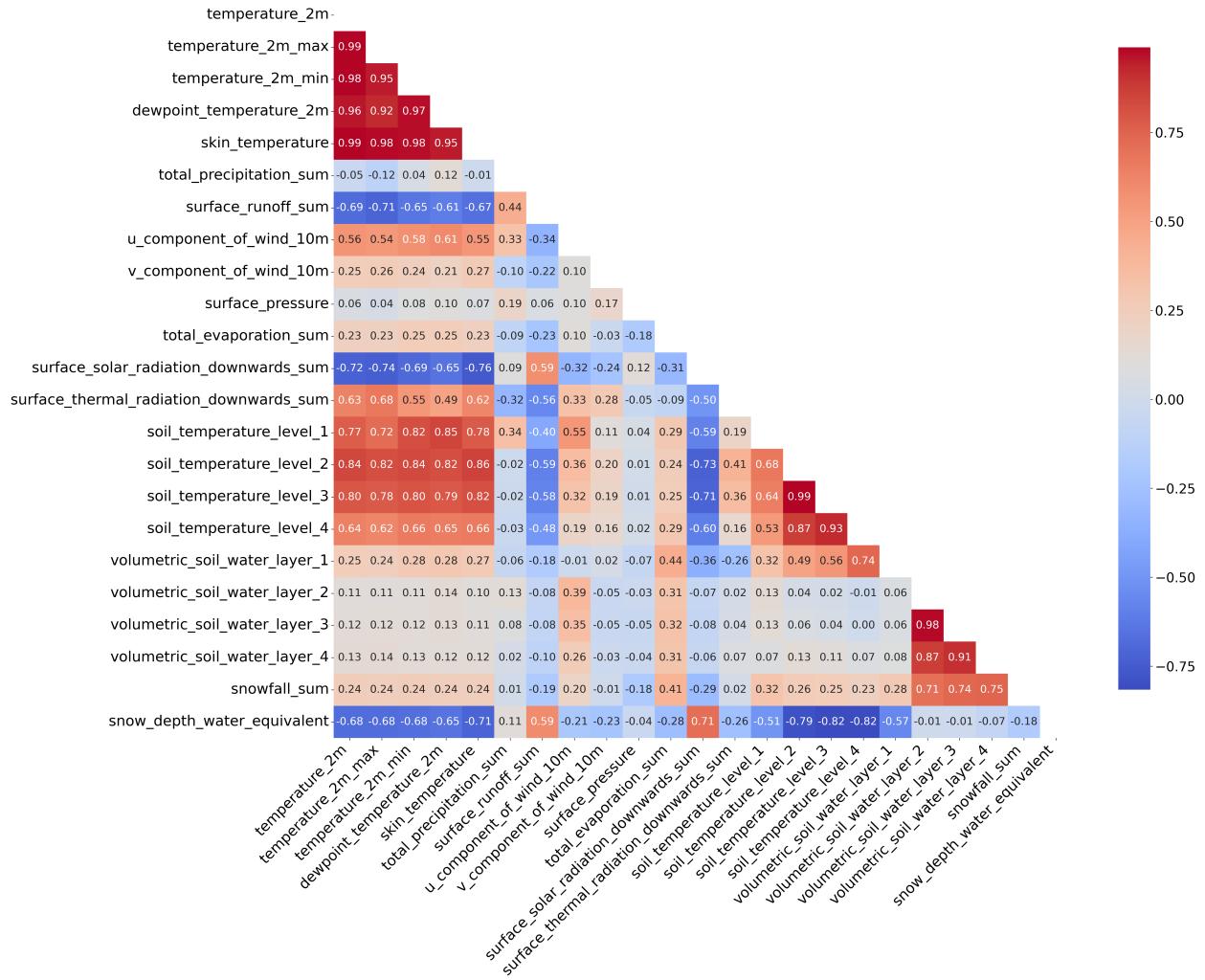


Figure S2: Correlation matrix of ERA5 meteorological variables showing multicollinearity among temperature-related fields and soil layers.

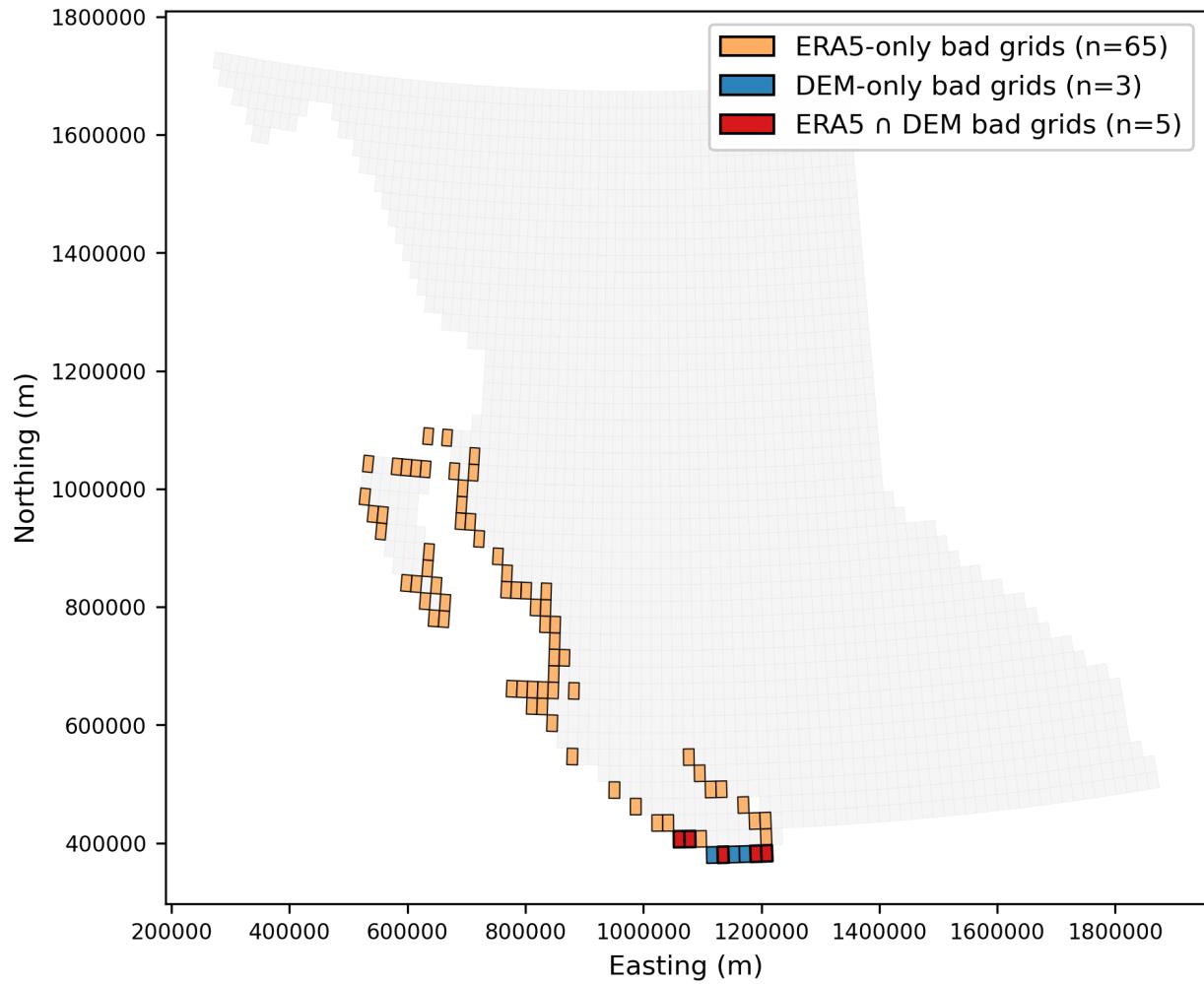


Figure S3: Spatial distribution of grids with missing ERA5 and/or DEM covariates across British Columbia.

Table S4: Summary of derived spatiotemporal features constructed for wildfire modelling.

Feature name	Type	Description
temp_lag1, temp_lag3, temp_lag7	Numeric	Lagged 2 m air temperature at 1-, 3-, and 7-day intervals
temp_rol3, temp_rol7, temp_rol30	Numeric	Rolling mean temperature over 3-, 7-, and 30-day windows
precip_sum7, precip_sum30, precip_sum60	Numeric	Rolling sums of total precipitation over 7-, 30-, and 60-day windows
consecutive_dry_days	Numeric	Number of consecutive days with negligible precipitation
wind_speed	Numeric	Computed daily wind speed magnitude
wind_max7	Numeric	Maximum wind speed over a 7-day window
wind_dir_sin, wind_dir_cos	Cyclic	Sine and cosine encoding of wind direction
month_sin, month_cos	Cyclic	Cyclical encoding of calendar month
is_fire_season	Binary	Indicator denoting the core wildfire season (April–September)
days_since_last_fire	Numeric	Number of days since the most recent fire occurrence within each grid cell.
historical_fire_count	Numeric	Cumulative count of historical fire occurrences for each grid cell.
historical_fire_prob_month	Numeric	Grid- and month-specific historical fire probability estimated
elevation_temp_interaction	Numeric	Interaction between elevation and air temperature
slope_wind_interaction	Numeric	Interaction between terrain slope and wind speed
heatwave	Binary	Indicator of extreme heat events based on high temperature thresholds.
extreme_wind	Binary	Indicator of extreme wind conditions based on high wind speed quantiles.
extreme_drought	Binary	Indicator of prolonged drought based on extended dry-day duration.
fuel_load	Numeric	Vegetation fuel load proxy, constructed as NDVI multiplied by consecutive dry days.
vegetation_moisture_stress	Numeric	Vegetation moisture stress index, defined as the ratio of NDVI to near-surface soil moisture.

Table S5: Confusion matrices for four model configurations on the test set.

Standard		Seq21			
	Pred 0	Pred 1			
Actual 0	777,134	12,906	Actual 0	785,728	4,312
Actual 1	1,039	351	Actual 1	1,228	162
Seq7		Reduced			
	Pred 0	Pred 1		Pred 0	Pred 1
Actual 0	776,914	13,126	Actual 0	780,586	9,454
Actual 1	1,015	375	Actual 1	1,099	291

Table S6: Seasonal classification performance of wildfire prediction by model configuration.

Fire Season									
Model	Accuracy	Precision	Recall	F1-score	TN	FP	FN	TP	
Standard	0.9670	0.0265	0.2696	0.0482	405,599	12,901	951	351	
Seq21	0.9881	0.0404	0.1244	0.0610	414,653	3,847	1,140	162	
Seq7	0.9665	0.0278	0.2880	0.0507	405,374	13,126	927	375	
Reduced	0.9751	0.0299	0.2235	0.0527	409,052	9,448	1,011	291	
Non-Fire Season									
Model	Accuracy	Precision	Recall	F1-score	TN	FP	FN	TP	
Standard	0.9998	0.0000	0.0000	0.0000	371,535	5	88	0	
Seq21	0.9985	0.0000	0.0000	0.0000	371,075	465	88	0	
Seq7	0.9998	0.0000	0.0000	0.0000	371,540	0	88	0	
Reduced	0.9997	0.0000	0.0000	0.0000	371,534	6	88	0	