

This manuscript presents WSDS-CAN, a multi-layer, machine-learning-ready wildfire dataset for the Canadian boreal forest that integrates final burned-area perimeters from the Canadian National Fire Database (CNFDB) with topographic, fuel, and meteorological covariates, packaged into fixed 128×128 tensors at an event-adaptive spatial resolution. High-resolution and spatially accurate fire perimeter datasets are valuable for the study of fire regimes and for spread modelling, and the dataset addresses a genuine gap by including small fires at an adaptive resolution for the Canadian landscape. However, the manuscript exhibits several structural, methodological, and data-quality issues that must be addressed to improve its clarity, rigour, and scientific contribution. A major revision is therefore recommended for the manuscript, with detailed comments provided below.

General Comments

1. Section 2.1.1 states that the ignition layer is a normalized Gaussian distribution whose pixel values represent a probability in $[0, 1]$. However, nearly one third of the 2,565 ignition rasters are numerically corrupted and violate this description. For example, 15HB-CASE_ignition.tif contains Inf, 15PA-ATIR02_ignition.tif contains NaN, and the maximum of 15WY-RIVER_ignition.tif is 2.37×10^{38} . This breaks the loss design in Section 3, which assigns a weight of $w = 50$ to ignition pixels with values exceeding 0.4, a criterion that is undefined when the values are greater than 1, NaN, or Inf. Please regenerate the layer to guarantee a $[0, 1]$ range for all ignition rasters.
2. The spatial coverage is far narrower than claimed, and the exclusions are largely unexplained. WSDS-CAN provides data for only four provinces, namely Manitoba, Saskatchewan, Ontario, and Quebec, and omits important boreal forest regions such as Yukon and the Northwest Territories. The authors should include the eligible fire events from the remaining provinces and territories and explain the reasons for excluding the other fire events. A map showing the distribution of the fire events contained in WSDS-CAN should also be added to the manuscript.

3. The temporal coverage of WSDS-CAN is not 2001 to 2020. Quebec is restricted to 2018 to 2020, Manitoba to 2014 to 2020, Saskatchewan to 2006 to 2020, and Ontario to 2009 to 2020. MODIS land cover is available for 2001 to 2024, so there is no auxiliary-data reason to omit those years, contrary to the statement that the record was restricted to 2001 to 2020 because of land-cover constraints. Please restate the true temporal span and discuss the resulting inter-annual and regional sampling bias for machine learning training. The CNFDB provides fire events up to 2024; if feasible, the spatial extent of WSDS-CAN could be expanded to include all boreal forest fires with a burned area greater than or equal to 0.01 km² across all Canadian provinces and territories, and the temporal range extended to cover 2001 to 2024.

4. WSDS-CAN provides only the final fire perimeter rather than daily (or any other temporal frequency) spread results. The increasing per-day IoU in Figure 5 reflects convergence toward the final mask, not validation against true daily fire fronts. Please either (a) rename the task to final perimeter prediction, or (b) derive day-of-burning perimeters from VIIRS or MODIS active-fire detections, because the CNFDB perimeters alone cannot support a spread task.

5. The covariate set is thin for the Canadian context. Compared with the 50 covariates of the Canadian Fire Spread Dataset (CFSD), WSDS-CAN provides only the u and v wind components, skin temperature, 2-m dewpoint temperature, elevation, slope, aspect, land cover, and leaf area index. The Fire Weather Index components and precipitation, which are standard drivers of Canadian fire behaviour, are not included. In addition, the manuscript uses ERA5-Land daily post-processed statistics but never states whether these are daily means or daily maxima. Please add further covariates and state the daily-statistic convention explicitly.

6. WSDS-CAN is described as a dataset for the boreal forest, but the authors do not explain how they retained only boreal forest fires and removed fires of other land-cover types. For example, the dominant land-cover type of ON_AYL1_2010 is Urban and Built-up Lands, and QC_20181080442 is

dominated by Croplands. Applying a fuel-based filter or, at a minimum, reporting the land-cover composition of the whole dataset are recommended.

Specific Comments

1. Table 2 lists only a summary of the raw source data. Please add a summary of all variables provided in WSDS-CAN.
2. Section 2.1.1 states that the Gaussian blob is centered on the reported ignition coordinates in one place and that it is maximized at the centroid in another. Please state which convention is used.
3. Section 3 states that land cover is one-hot encoded into eight channels, but the data contain 17 classes, and the 17-to-8 mapping is unspecified. The IoU values in Table 3 (approximately 0.48 to 0.51) are modest and are not compared with any external baseline. Please provide the class mapping and the training details.
4. Lines 174 to 176 indicate that only two unique LAI frames were generated for fires lasting fewer than five days. However, for the fire event 08LA-HAWKINS, which lasts one day, all LAI files are identical. Moreover, even for fire events lasting more than five days, for example MB_NE093_2015 (54 days), MB_NE098_2015 (54 days), and MB_NE110_2015 (52 days), the five LAI frames are also completely identical and provide no temporal variation of LAI. Please correct all files affected by this problem.
5. The authors should explicitly state in Section 2.1 which programming language was used for data preprocessing and model implementation.
6. In Figure 2, the historical fire data include Start_date and End_date, but the WSDS-CAN output does not provide the precise daily start and end dates for each fire.
7. The mask values of MB_EA050_2015 and MB_NE027_2014 range from 0 to 255 rather than from 0 to 1. Please resolve this problem.

8. All four meteorological variables of MB_NE073_2017 are empty. Please check all files, repair or remove the affected events, and state the quality-control criterion.
9. Lines 179 to 181: The use of the leaf area index as a proxy for fire progression is weakly supported. The MODIS LAI product (8-day, 500 m) is too coarse and varies too slowly to capture fire advance over an event lasting a few days, and it is contaminated by smoke and cloud during active burning. The claim that low-LAI pixels correlate with fire growth across the five frames is not well supported, particularly given that many events contain identical LAI frames. This interpretation should be tempered.
10. Line 204: The demonstration uses a random 70:15:15 split. For wildfire spread modelling this risks spatial and temporal leakage, because neighbouring fires from the same region and year may be divided between the training and test sets, which inflates the reported metrics. Spatial or temporal block splitting is the standard practice and should be adopted or at least discussed.
11. The fire_id naming scheme is not uniform. For example, some Manitoba identifiers take the form MB_NE073_2017 while others take the form MB_MB2018NE116_2018 with a duplicated province code, and Saskatchewan identifiers such as 15WY-LRIR02 carry no province prefix whereas the Manitoba, Ontario, and Quebec identifiers do. A uniform identifier convention should be adopted.

Technical Corrections

1. Line 46: NDWS (Huot et al., 2022) reports 18,545 fire events; the manuscript states 18,845. Please correct.
2. Table 2: "EAR5" should be "ERA5-Land".
3. Line 184: ERA5-Land is provided on a regular 0.1-degree geographic grid (EPSG:4326), not a Sinusoidal projection. Please correct the projection statement.
4. Line 221: "unite weight" should be "unit weight".

5. Figure 3 caption: "ignition" should be "Ignition", "file" should be "fire", and "Km²" should be "km²".
6. Table 1: "CFSDS" should be "CFSD".
7. "Temperature of the surface of the Eart" should read "Earth".
8. Line 60: "specially" should be "especially".
9. Figure 1 caption: in "Distribution of fire events greater than 1 ha during 2001 and 2020 across Canada", "during" should be "between".
10. Table 2: the unit of the leaf area index should be m²/m².