General assessment

This study develops a high-resolution, machine-learning-ready wildfire and environmental dataset for Canada. The dataset could serve as a valuable foundational resource for advancing wildfire prediction and real-time monitoring in northern forest environments. The topic is important and the overall methodological framework appears sound. However, several aspects of the manuscript require further clarification and reorganization before the contribution and utility of the dataset can be fully assessed.

Major comments

1. **Strengthen the motivation for focusing on Canada.**
   I suggest adding more background in both the Abstract and Introduction to emphasize the global significance of Canadian wildfires and to better justify Canada as the study region. For example, why are Canadian wildfires particularly important and urgent to study? What major ecological, climatic, socioeconomic, or air-quality impacts have recent Canadian wildfire events caused? A stronger motivation would help clarify the significance of the dataset and its broader relevance.
2. **Clarify and, if possible, extend the study period.**
   The dataset covers 2001–2020. One justification provided around Line 75 is that "due to historical constraints in auxiliary terrain data (e.g., land cover), the dataset was restricted to events occurring between 2001 and 2020." However, land-cover products such as MODIS MCD12Q1 are available beyond 2020. I therefore suggest either extending the dataset to more recent years or providing a clearer explanation of why 2020 was selected as the endpoint.
3. **Reorganize Section 2 for clarity.**
   Section 2 is currently difficult to follow and contains some repetitive information. I suggest reorganizing this section to align more closely with Figure 2. Specifically, consider adding a brief overview paragraph at the beginning of Section 2, followed by four subsections describing the datasets and preprocessing procedures for:
   2.1 Fire data
   2.2 Fuel data
   2.3 Topographic data
   2.4 Meteorological data
   This structure would make the data-processing workflow much easier to follow.
4. **Discuss transferability of the framework.**
   Please discuss whether this framework can be transferred or adapted to other regions globally. This would help demonstrate the broader utility of the dataset-development approach beyond Canada. It would also strengthen the reproducibility of the study if the processing scripts could be shared together with the dataset on Zenodo.

Minor comments

- **Line 74:** It is not clear where the land-cover data come from. Why are land cover data categorized as terrain data here, while they are treated as fuel data in Figure 2? Please clarify the source of both the terrain and land-cover datasets. It may also be helpful to introduce these datasets before referring to them in this sentence.

- **Lines 74–75:** Consider moving the description of the study period to the end of the data section and explaining that the selected period represents the common temporal overlap among all datasets.
- **Line 75:** Please provide a justification for constraining wildfire events to durations between 1 and 60 days.
- **Line 80:** The manuscript states that "we aggregated these terrain and meteorological data at the provincial and territorial levels." If the variables were aggregated to such coarse administrative units, how is the "high-resolution" nature of the dataset maintained? Please clarify the spatial resolution used in the final dataset.
- **Line 85:** Please define what is meant by "secondary spatial variables."
- **Line 90:** Instead of "to incorporate a proxy for vegetation density and biomass," would "to represent fuel amount" or "to characterize fuel quantity" be more direct and consistent with the purpose of the variable?
- **Line 100:** Several meteorological variables are included. Please provide additional justification for why these particular variables were selected and explain their relevance to wildfire behavior or prediction.
- **Table 2:** I suggest adding a new leftmost column to group the variables into categories consistent with Figure 2, such as fire, fuel, topography, and meteorology. I also suggest including the fire-related variables in this table. The caption could be revised to: **"Summary of variables and their data sources used to create WSDS-CAN."**
- **Line 110:** Please clarify the source of the "ignition points." If they are obtained from CNFDB, this should be introduced earlier when the fire dataset is first described.
- **Figure 2:**
    1. Remove "historical" before "fire data," because all datasets shown in the figure are historical rather than only the fire data.
    2. Add the data source for each dataset directly in the figure or caption, for example, fire data from CNFDB.
    3. Explain in the caption what the reported number of layers represents.
- **Figure 3 caption:** Change "ignition" to "Ignition."
- **Figure 4:** Please add a legend for LAI.
- **Line 195:** Please check the formatting/typo in "128 × 128-pixel."