International Journal of Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: www.tandfonline.com/journals/tres20

Burned area mapping across the Arctic-boreal zone with Landsat and Sentinel-2 imagery

Stefano Potter, Yili Yang, Arden Burrell, Anna Talucci, Andrew A. Clelland, Sander Veraverbeke, James T. Randerson, Scott J. Goetz, Logan T. Berner, Michelle C. Mack, Xanthe J. Walker, Susan M. Natali & Brendan M. Rogers

To cite this article: Stefano Potter, Yili Yang, Arden Burrell, Anna Talucci, Andrew A. Clelland, Sander Veraverbeke, James T. Randerson, Scott J. Goetz, Logan T. Berner, Michelle C. Mack, Xanthe J. Walker, Susan M. Natali & Brendan M. Rogers (05 Mar 2026): Burned area mapping across the Arctic-boreal zone with Landsat and Sentinel-2 imagery, International Journal of Remote Sensing, DOI: 10.1080/01431161.2026.2639127

To link to this article: https://doi.org/10.1080/01431161.2026.2639127

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

View supplementary material

Published online: 05 Mar 2026.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tres20

INTERNATIONAL JOURNAL OF REMOTE SENSING
https://doi.org/10.1080/01431161.2026.2639127

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Burned area mapping across the Arctic-boreal zone with Landsat and Sentinel-2 imagery

Stefano Potter[a], Yili Yang[a], Arden Burrell[a], Anna Talucci[a], Andrew A. Clelland[b,c], Sander Veraverbeke[d,e], James T. Randerson[f], Scott J. Goetz[g], Logan T. Berner[g], Michelle C. Mack[h], Xanthe J. Walker[h], Susan M. Natali[a] and Brendan M. Rogers[a]

[a]Woodwell Climate Research Center, Falmouth, MA, USA; [b]Atmosphere Ice and Climate, British Antarctic Survey, Cambridge, UK; [c]Department of Computing, Imperial College London, London, UK; [d]School of Environmental Sciences, University of East Anglia, Research Park, Norwich, UK; [e]Department of Earth Sciences, Vrije Universiteit, Amsterdam, The Netherlands; [f]School of Physical Sciences, University of California, Irvine, CA, USA; [g]Computing and Cyber Systems, Northern Arizona University, Flagstaff, AZ, USA; [h]Biological sciences, Northern Arizona University, Flagstaff, AZ, USA

ABSTRACT

Wildfires in the Arctic-boreal zone have increased in frequency over recent decades, carrying substantial ecological, social, and economic consequences. Remote sensing is crucial for mapping burned areas, monitoring wildfire dynamics, and evaluating their impacts. However, existing high-latitude burned area products suffer from significant discrepancies, particularly in Siberia, and their coarse spatial resolutions limit accuracy and utility. To address these gaps, we developed a convolutional neural network model to map burned areas at a 30 m resolution across the Arctic- boreal zone using Landsat and Sentinel-2 imagery. Using vegetation indices including the normalized burn ratio, normalized difference vegetation index, and normalized difference infrared index our model achieved strong performance, with an Intersection Over Union (IOU) of 0.77 and an F1 score of 0.85 on unseen test data. Performance was higher in North America (IOU = 0.84) than in Eurasia (IOU = 0.72), reflecting regional differences in fire regimes and data quality. Predictions for six representative years showed our model's burned area closely matched the median values of Landsat, MODIS, and VIIRS-based products, although alignment varied annually and spatially. Visual assessments indicated our approach was generally more accurate, notably in detecting unburned vegetation islands within fire perimeters missed by other products. This research has numerous potential applications, such as analysing feedback between vegetation and burn patterns, characterizing spatial dynamics of unburned islands, and improving carbon emission estimates through detailed burn severity assessments.

ARTICLE HISTORY
Received 16 October 2025
Accepted 24 February 2026

KEYWORDS
Wildfire; remote sensing; deep learning; Arctic

CONTACT Stefano Potter spotter@woodwellclimate.org Woodwell Climate Research Center, 149 Woods Hole Road, Falmouth, MA 02540, USA

Supplemental data for this article can be accessed online at https://doi.org/10.1080/01431161.2026.2639127

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

1. Introduction

The Arctic-boreal zone (ABZ) is a vast, high-latitude region that encircles Earth's northernmost land masses. It includes areas of boreal forest, which is the world's largest land biome (Gauthier et al. 2015), and tundra, which is typically treeless, cold and underlain with permafrost (Elmendorf et al. 2012). The ABZ is warming two to four times faster than the global average (Rantanen et al. 2022; Serreze and Barry 2011), and subsequently both boreal (Kelly et al. 2013) and tundra (ACIA 2004) ecosystems have experienced increasing fire frequency in recent decades (French et al. 2015; Jones et al. 2024). Shifts in future fire regimes will likely alter ecosystem dynamics, including vegetation composition (Flannigan et al. 2009; Johnson and Johnson 1996; Johnstone et al. 2010; Mekonnen et al. 2019; Payette 1992; Viereck 1973) and carbon and nutrient cycling (Balshi et al. 2009; Genet et al. 2013). These northern ecosystems store about 50% of global belowground soil carbon and play a crucial role in global climate regulation (McGuire et al. 2009; Young et al. 2017). However, the increasing frequency of fires has the potential to shift these ecosystems from historical carbon sinks to sources (Bond-Lamberty et al. 2007; Li, Lawrence, and Bond-Lamberty 2017; Virkkala et al. 2025; Walker et al. 2019; Wang et al. 2021). To understand how changing fire regimes influence carbon and other ecosystem dynamics, it is critical to accurately map ABZ burned areas.

Fires in the North American boreal typically occur in mature stands of black spruce (*Picea mariana*) and jack pine (*Pinus banksiana*), both of which are prone to crown fires that kill most of the standing trees (de Groot et al. 2013; Johnstone et al. 2010; Rogers et al. 2015; Scherer-Lorenzen and Schulze 2005). Such crown fires exert large changes in optical to shortwave infrared reflectance, and hence are well suited to being mapped from satellite imagery. Burned area mapping in Alaska historically relied on digitized maps of fire observations from the Alaska Large Fire Database (Kasischke, Williams, and Barry 2002) and the Monitoring Trends in Burned Severity program (MTBS; Eidenshink et al. 2007), which is a satellite-based mapping project that relies largely on human interpreters for imagery selection and mapping. For instance, a human interpreter examines pre-fire and post-fire scenes and then creates a normalized burn ratio (NBR) image that is used to delineate a burned area boundary. In Canada, burned area mapping historically relied on the Canadian National Fire Database (CNFD; Amiro et al. 2001; Stocks et al. 2002), which primarily maps perimeters using manual digitization and aerial Global Positioning System (GPS). More recently, the Canadian National Burned Area Composite (NBAC; Hall et al. 2020; Skakun et al. 2022) was created and improves on the CNFD by using remote sensing imagery to account for water bodies and unburned islands within the fire perimeters.

Both MTBS (1984–2024) and NBAC (1986–2024) offer relatively long time series of burned area. However, these products are not fully consistent with one another, primarily because MTBS does not account for unburned vegetation patches within its vector-based fire perimeter polygons. Additionally, these datasets are limited geographically to Alaska and Canada, respectively, and comparable products do not currently exist for Eurasia.

Boreal Eurasian forests are dominated primarily by Scots Pine (*Pinus sylvestris*) and several species of larch (*Larix* spp.) (Furyaev et al. 2001; Goetz et al. 2007; McRae et al. 2006; Rogers et al. 2015; Scherer-Lorenzen and Schulze 2005), which are more prone to lower intensity and non-stand replacing surface fires than boreal North America forests. The spectral change from surface fires is typically much smaller (<100 ha) and short-lived

compared to crown fires, and hence boreal Eurasian fires are often less amenable to satellite-based detection. Additionally, no reliable government burned area mapping data source exists in Siberia, and official statistics largely underestimate burned area (Rogers et al. 2015; Romanov et al. 2022; Shvidenko and Schepaschenko 2013; Shvidenko et al. 2011). While the total burned area in Eurasia does not occur solely in Siberia, the large majority of it it does (Rogers et al. 2015). However, substantial discrepancies persist among these products in estimating burned area across Eurasia, especially in southern Siberia (Clelland et al. 2024). This region predominantly consists of open forests, grasslands, and shrublands, which burn early in the season, mostly by small fires (Scholten et al. 2024). These fires are particularly challenging to detect via satellite remote sensing due to the rapid vegetation recovery. Given these inconsistencies in Eurasia, and the lack of reliable government data, it is important to develop an independent and accurate burned area mapping approach that can be uniformly applied to the ABZ at high (30 m) resolution.

Global burned area mapping from 2001 to present has historically been derived from 250 to 500 metre (m) Moderate Resolution Imaging Spectroradiometer (MODIS) and Visible Infrared Imaging Radiometer Suite (VIIRS) products, such VNP64A1 (Giglio 2019), MCD64A1 (Giglio et al. 2018), Fire Climate Change Initiative (FireCCI) (Chuvieco et al. 2016, 2018) and Arctic Boreal Burned Area (ABBA; Loboda et al. 2024). MODIS sensors have provided frequent acquisitions, which are favourable to burned area mapping, but there are known omission errors as small (<100 ha) fires are missed (Y. Chen et al. 2023). MODIS sensors also tend to overestimate burned area at the pixel level due to their relatively coarse resolution, which is prone to including unburned vegetation islands and water bodies (Giglio et al. 2018). More recently the use of 30 m Landsat mostly bypasses these issues because of its higher spatial resolution (Guindon et al. 2018; Walker et al. 2018), but it suffers from less frequent overpass times (16-day return intervals) that are compounded by often cloudy conditions in the boreal and tundra biomes, leading to data gaps. More recently, the launches of Sentinel-2A (launched 2015) and 2B (launched 2017) began a new era of higher-resolution satellite monitoring (Drusch et al. 2012; Roy et al. 2014). The return interval of both Sentinel-2 satellites is 5 days, but when combined with Landsat 8 this is further reduced to 2.9 days, which is a large improvement compared to the 16-day Landsat return interval (Roy et al. 2019).

Satellite-based burned area mapping often makes use of vegetation indices (VIs) such as NBR, which uses the near-infrared (NIR) and shortwave infrared (SWIR) reflectance bands (Key and Benson 2006):

$$NBR = \frac{\mathrm{NIR} - \mathrm{SWIR}_2}{\mathrm{NIR} + \mathrm{SWIR}_2} \tag{1}$$

NIR is sensitive to the leaf structure and area of live vegetation, and SWIR is sensitive to water content in soils and vegetation. Additionally, SWIR is useful for separating dead wood from soil and ash (Jia et al. 2006). Both Landsat and Sentinel-2 offer two SWIR bands, with SWIR_1 centered at 1.61 μ m and SWIR_2 at 2.20 μ m. The differenced normalised burn ratio (dNBR), which is the difference between pre- and post-fire NBR, has become standard for mapping using spectral indices (Eidenshink et al. 2007; Fassnacht et al. 2021). The dNBR is often combined with a threshold based on ground observations to separate burned and unburned pixels (Loboda, O'neal, and Csiszar 2007; Rogers et al.

2014; Veraverbeke, Rogers, and Randerson 2015; Walker et al. 2018; Potter et al. 2023). A drawback of this threshold-based approach is that dNBR thresholds are often region-, vegetation-, and even fire-specific, and have to be fine-tuned using manual rule-based approaches (Key and Benson 2006; Loboda, O'neal, and Csiszar 2007; Miller and Thode 2007). Additionally, since the threshold does not take into account the spatial surroundings of a pixel, it can confuse burned pixels with other plot-level disturbances, for example, due to changing hydrology, insects, pathogens, land use, extreme weather events, or abrupt permafrost thaw, especially at 30 m (Potter et al. 2023).

In addition to improvements in spatial, temporal and radiometric quality of data, there have also been advancements in machine and deep learning, particularly with respect to image classification and image segmentation, which has led to improvements in burned area mapping (Jain et al. 2020). Automated machine learning methods such as support vector machines (Cao et al. 2009; Pereira et al. 2017), random forests (Ramo and Chuvieco 2017; Roteta et al. 2019) and boosted regression trees have shown promise in burned area mapping due to their ability to model complex non-linear phenomena. For example, Todd et al. (2017) used MTBS data, boosted regression trees and Landsat imagery to map burned area in the contiguous United States from 1984 to 2015, finding that this method mapped 36% more burned area than the Global Fire Emission Database version 4 (GFED4; Giglio, Randerson, and Van Der Werf 2013; Randerson et al. 2012; Van der Werf et al. 2010), and 116% more than MTBS.

Long et al. (2019) mapped global burned area at 30 m using a Random Forest model with eight VIs and six surface reflection bands derived from Landsat 8 imagery. The authors used pre- and post-fire image differencing with the VIs to achieve a global commission error of 13% and omission error of 30%. Accuracy ranged from 71% to 96% depending on the region tested. While validation scenes were located in Alaska, Canada and Eurasia, no error metrics specifically for these regions were reported. Additionally, validation was only reported for the 2015 fire year, and there are known issues in the burned area estimations in other years, partly related to the Landsat 7 ETM+ Scan Line Corrector (Clelland et al. 2024).

Deep learning, and more specifically Convolutional Neural Networks (CNNs), have great potential for mapping burned areas at high resolution because it is designed to operate on spatially dependent data, such as images, and can effectively capture both local and global dependencies. CNNs excel at detecting spatial patterns through convolutional layers (LeCun et al. 1998) and apply nonlinear functions that generalize well to learned features (Knopp et al. 2020). For example, Hu, Zhang, and Ban (2023) tested a variety of UNet CNN models using MTBS training polygons and Landsat imagery in the contiguous U.S. from 2001 to 2019 to map burn severity, and found that an attention UNet achieved the best mean Intersection Over Union (IOU) score of 0.78 and a Kappa score $\approx$ 0.90.

Most recently Zhang et al. (2024) examined various UNet architectures along with multi-source satellite imagery including Synthetic Aperture Radar (SAR) from Sentinel-1 and surface reflectance from Sentinel-2 to map burned area at 20 m in Canada. NBAC was used as the reference fire perimeters to train the model over the 2017 and 2018 fire seasons, and models were evaluated on the 2019 fire season. They found that UNets worked best for single sensor data, such as using only SAR or only surface reflectance rather than both together. They also tested if

pre- and post-fire image differencing performed better than only post-fire scenes, and found that pre- and post-fire imagery with surface reflectance alone resulted in the best model with an IOU ≈ 0.84.

Despite these advances, very few research efforts have attempted to map wildfires across the whole ABZ at 30 m, where forests and permafrost dynamics are strongly influenced by wildfire activity with important implications for global carbon budgets (Phillips et al. 2022; Walker et al. 2018, 2019). One notable exception is the Global Annual Burned Area Maps (GABAM; Long et al. 2019), although known issues with false positives at high latitudes have been reported (Clelland et al. 2024). In this work, we introduce new methods for mapping burned areas at 30 m spatial resolution across the ABZ. We integrate Landsat 5, 7, 8, and 9 imagery with Sentinel-2A and 2B data to train two models: a machine-learning-based extreme Gradient Boosting (XGBoost) model and a deep-learning-based UNet++, which is a type of CNN. Our overarching goals in this research were to 1) evaluate the difference in model performance between machine learning and deep learning methods, 2) determine if region-based models performed better or worse than a model trained across the ABZ, and 3) assess burned area estimates, both quantitatively, for representative fire years, and qualitatively through visual comparisons of individual fire events. Our objective is to produce an annual burned area mapping product spanning 2001–2024, which will be updated annually using our methodology. Here, we evaluate the effectiveness of our methods using six representative fire years from the ABZ.

2. Materials and methods

We compiled a large reference dataset of fire perimeters across North America and Siberia to train, validate, and test our models. For each fire, we used Google Earth Engine to acquire and preprocess Landsat and Sentinel-2 surface reflectance imagery, removed clouds and snow, calculated spectral indices, and harmonized bands across sensors. We then created two types of image composites including annual and multi-monthly to capture pre- and post-fire conditions, and generated predictor layers representing spectral change. These composites, along with rasterized fire perimeters, were used to train XGBoost and U-Net++ models under multiple sampling and cross-validation schemes. Model performance was evaluated using a suite of accuracy metrics and tested across different ecoregions. Finally, we applied our best-performing models across the ABZ, constrained predictions using active fire detections, and compared results against ten existing burned-area products. A workflow of our methodology is shown in Figure 1.

2.1. Study area

Our study area (Figure 2) is the ABZ, defined by boreal and tundra biomes according to Dinerstein et al. (2017). It includes areas within Alaska, Canada, Greenland, Iceland, Fennoscandia, and Russia. Altogether, the study area is 25,462,143 km2, representing 17% of Earth's land surface.

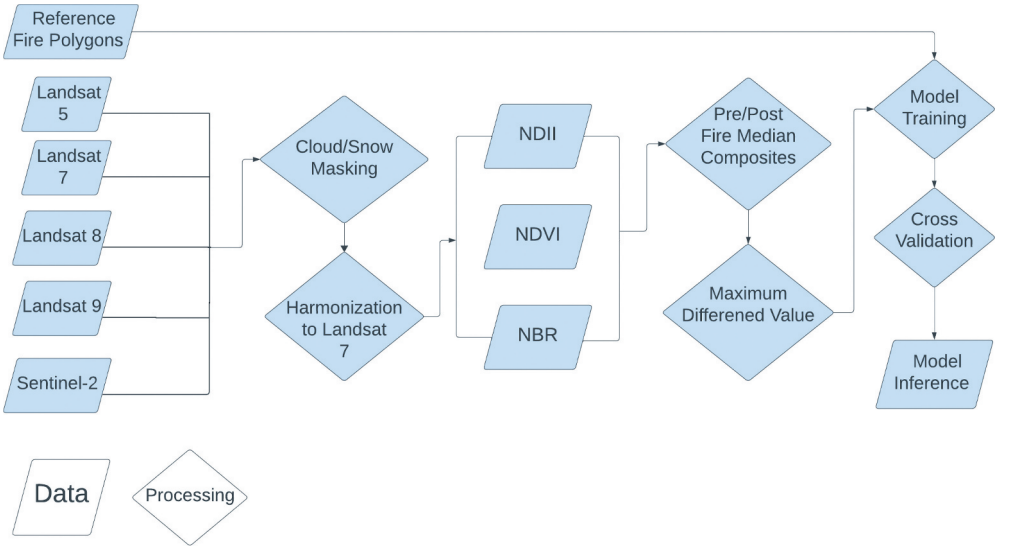

Figure 1. Simplified flowchart of methodology. Acronyms include normalised difference infrared index (NDII), normalised difference vegetation index (NDVI) and normalised burn ratio (NBR).

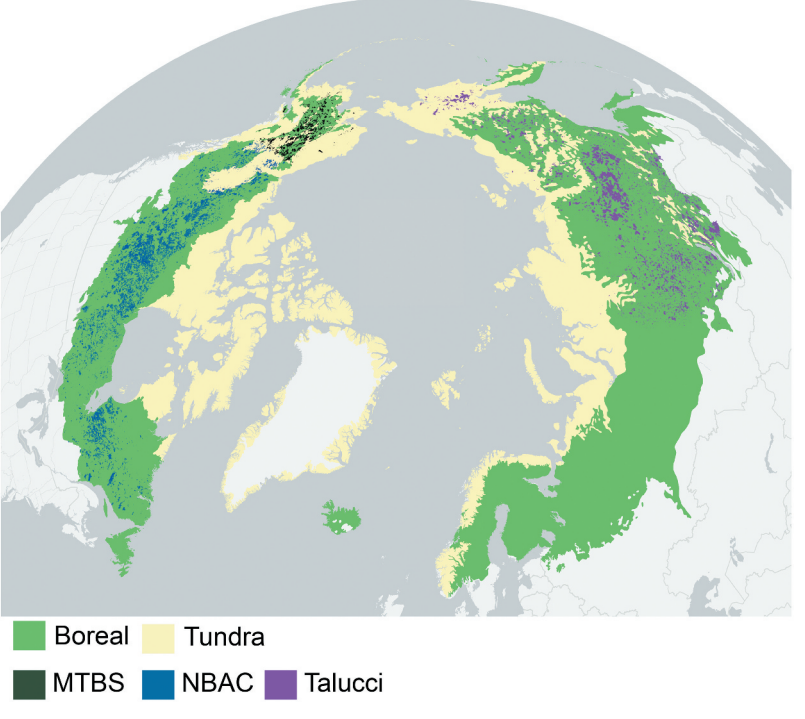

Figure 2. Map of the study area. Historical fire polygons within the boreal and tundra regions used for model training, validation and testing are shown in black for MTBS in Alaska, in blue for NBAC in Canada, and in purple for Talucci, Loranty, and Alexander (2022) in Russia.

2.2. Reference fire polygons

For model training, validation and testing we used a total of 12,425 reference fire polygons covering 84.53 Megahectares (Mha) in North America. Fire polygons in Canada ($n = 11{,}248$, burned area = 70.74 Mha) were acquired from NBAC, and polygons in Alaska ($n = 2{,}076$, burned area = 9.59 Mha) were acquired from MTBS (Figure 1). We include NBAC records from 1986 to 2020 and MTBS records from 1984 to 2020.

In Eurasia, our reference polygons were limited to eight ecozones across eastern Siberian boreal and tundra regions (Talucci, Loranty, and Alexander 2022). Fire polygons are available in this region from 2001 to 2020 and were developed by creating 30 m Landsat image composites and calculating the associated dNBR, which was then converted to binary burned and unburned pixels based on dNBR thresholds (Key and Benson 2006). To delineate fire perimeters from binary burned/unburned pixels, MODIS hotspot data were used. Specifically, a 3 km buffer was applied around MODIS hotspot points, merged into single polygons and fire perimeters were delineated within these larger polygons due to computational limits. For more detailed discussion of these methods see Talucci, Loranty, and Alexander (2022). A total of 7740 fire polygons covering 76.84 Mha in Siberia were used in model training and validation (Figure 1). Although this reference dataset does not encompass the entirety of Eurasia, 75% of the historical burned area between 2001 and 2023 occurred in the region containing the Talucci, Loranty, and Alexander (2022) reference polygons according to MCD64A1. Table S1 compares the three reference polygons sources, their temporal availability, and reported accuracy metrics if available.

After acquiring the vector fire perimeters from the respective sources, we uploaded the data to Google Earth Engine (GEE; Gorelick et al. 2017). We then calculated the bounding box of every fire perimeter, buffered them by 10 km, and used this buffered bounding box as the area of interest for every individual fire. We buffered the bounding box for two reasons: 1) to ensure we trained on both burned areas and unburned areas outside of where the fire occurred, and 2) so that we could determine the influence of buffering distance from MODIS and VIIRS active fire hits on model performance (section 2.2). We then acquired available Landsat and Sentinel-2 imagery, also within GEE. We used Landsat 5, 7, 8 and 9 Collection 2 with Sentinel-2A and 2B surface reflectance for image acquisition. Pre-processing steps for Landsat imagery included removing clouds, cloud shadows, and snow using the C Function of Mask (CFMask; Foga et al. 2017).

We then calculated NBR and two additional VIs, including the normalized difference infrared index (NDII; Hardisky, Smart, and Klemas 1983):

$$\mathrm{NDII} = \frac{\mathrm{NIR} - \mathrm{SWIR}_1}{\mathrm{NIR} + \mathrm{SWIR}_1} \tag{2}$$

and the normalized difference vegetation index (NDVI; Tucker 1979):

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{3}$$

NDII is a proxy for vegetation moisture content and NDVI is a proxy for chlorophyll content or vegetation greenness. All spectral bands used to calculate the VIs and for each satellite are listed in Table S2. Due to differences in the sensors and spectral bandwidths

between Landsat satellites, it is necessary to cross-calibrate bands and spectral indices between sensors (Berner et al. 2020; Roy et al. 2016; Sulla-Menashe, Friedl, and Woodcock 2016). To do so, we adjusted Landsat 5, 8 and 9 to match Landsat 7 using the cross-calibration approach implemented by the LandsatTS package in R (Berner et al. 2023). Specifically, we cross-calibrated NBR, NDVI and NDII with second- or third-order polynomial regression models that were fitted and evaluated using measurements from a random sample of recently undisturbed grid cells ($n \approx 100{,}000$) across the boreal forest biome. These grid cells were identified, and the Landsat data were sampled, as part of a prior study (Berner and Goetz 2022). Cross-validation showed high model performance ($R^2 = 0.94$–0.97), with the mean absolute percentage bias decreasing from 3.4% (range = −5.3% to 6.6%) to 0.2% (range = −0.5% to 0.0%). While Berner et al. (2023) did not explicitly develop correction coefficients for Landsat 9, we implemented the same correction coefficients developed for Landsat 8 because the bandwidths for each are identical.

Clouds were masked from Sentinel-2 using the Sentinel-2 Cloud Probability based on the s2cloudless algorithm (Skakun et al. 2022). We used a cloud probability threshold of 50%, i.e. any pixel with a greater than 50% probability of being a cloud was masked. This threshold provides a practical balance between minimizing cloud interference in the imagery and retaining as much cloud-free data as possible, ensuring that high-probability cloud pixels are effectively excluded without excessively discarding valuable clear-sky pixels. After masking clouds, we resampled the 10 m and 20 m resolution pixels to 30 m to match Landsat imagery by using spatial averaging. As with the Landsat data, we calculated NBR, NDVI and NDII and cross-calibrated Sentinel-2 to match Landsat 7 using the same polynomial regression coefficients used with Landsat 8.

Using the burned year associated with each fire, we then acquired pre-fire and post-fire imagery in two ways: using 1) annual and 2) multi-monthly image compositing methods. We tested these two methods because compositing provides a better opportunity to detect spectral changes immediately following a fire, which is especially important for capturing surface fires whose spectral signatures may fade within a small number of months, such as understory forest fires or fires in herbaceous and shrub-dominated ecosystems. Since the timing of each fire within a given year is unknown beforehand, the multi-monthly method helps ensure maximal spectral change is captured, regardless of when during the season the fire occurs. For 1) we acquired pre-fire imagery as the median composite across all available images one year before fire in the summer months (June-August) and post-fire imagery as the median composite one-year post-fire in the summer months. We then differenced the pre- and post-fire imagery by subtracting the post-fire from the pre-fire image (Figure S2). For 2) we acquired pre-fire imagery in multi-monthly time intervals while allowing for one month overlap from March to November, and post-fire imagery that matched the same pre-fire months used but within the year of fire. For example, if the year of fire was 2004, pre-fire intervals would be 2003–03-01 to 2003–04-30, 2003–04-01 to 2003–05-31 and continuing to 2003–09-01 to 2003–10-31. The same multi-monthly intervals were used for post-fire imagery, but images were selected in 2004 (year of fire). Median composites in each of the multi-monthly pre- and post-fire intervals were then calculated. We used a one month overlap to increase the number of potential observations per pixel that would be available for input into the median composites. We also created median image composites using the same methods

as 1). We then differenced the pre-fire and post-fire imagery and then calculated the maximum pixel values across all eight of the differenced image composites. A flowchart of this method is in Figure S3.

For method 2, we applied an additional pre-processing step to remove snow, which was more problematic for the multi-monthly approach because it used both spring and fall imagery. Before calculating our median composites, we first calculated the normalized differenced snow index (NDSI):

$$NDSI = \frac{\text{Green} - \text{SWIR}_1}{\text{Green} + \text{SWIR}_1} \tag{4}$$

NDSI ranges from −1 to 1 with pixels closer to 1 having a higher probability of snow. We removed any pixels with values >–0:2 to be conservative in our snow masks, and to avoid inclusion of any snow covered pixels (Riggs, Hall, and Román 2015). Hereafter method 1 will be referred to as the annual composite method, and method 2 will be referred to as the multi-monthly composite method.

Lastly, for both the annual and multi-monthly composite methods, we masked water bodies using the Pekel et al. (2016) yearly water classification product derived from Landsat.

For both image composite methods, we exported the dNBR, dNDVI and dNDII bands, which were used as predictor variables in our XGBoost and UNet ++ models (Figure S4). Additionally, we rasterized and exported the reference fire perimeters as a binary fourth band that represented the target variable as either 0 (no fire) or 1 (fire).

After exporting individual images of each fire, we normalized the predictor bands in each image so pixel values were stretched from 0 to 1. We first calculated the 5% quantile as the minimum and 95% quantile as the maximum value across all available training data in each band, and then normalized each image using these minimum and maximum values. The quantiles were used to remove abnormally high and low values in the training data, and ensured all training images were normalized consistently.

2.3. Model development

We tested two primary models: XGBoost (T. Chen and Guestrin 2016) and UNet++ (Zhou et al. 2019). Due to the relatively large input size of our training data (100 gigabytes), when comparing the XGBoost model to the UNet++ we randomly selected 1000 polygons each from MTBS, NBAC and Talucci, Loranty, and Alexander (2022), and only used those polygons for model training, validation and testing. This represents 98% of the MTBS polygons, 10% of the NBAC polygons and 13% of the Talucci polygons. We used a randomized search grid within a 5-fold cross-validation framework to tune the XGBoost model parameters (supplemental methods 1.1, Table S3). For final XGBoost model evaluation we used 70% of the 3000 fires as training, 20% as validation and 10% as testing.

Due to the input requirements of the UNet++, we extracted 128×128 sub-matrices from the original fire imagery rather than processing the full-sized scenes, allowing for a direct performance comparison with our XGBoost model. We then trained our UNet++ model in an identical manner as to XGboost where we used the same sample of 3000 fires. We also used the same 70/20/10 split for our training/validation/testing sets to match the XGBoost model and to be able to directly compare the two.

We then added a second experiment which only focused on UNet++ model performance. For this experiment we no longer focused on a sample of 3000 fires like we did when comparing XGboost and UNet++, but instead we used all available reference polygons.

Model testing was implemented using a 5-fold cross validation. While we use Landsat/Sentinel-2 imagery dating back to 1985 to train and validate our models, we restricted our model testing to regions surrounding known active fire detections from MODIS (Giglio and Justice 2015a,2015b) and VIIRS (Schroeder et al. 2014), similar to many previous approaches (Potter et al. 2023; Talucci, Loranty, and Alexander 2022; Veraverbeke, Rogers, and Randerson 2015). By focusing our testing on these known active fire areas, we aim to minimize commission errors, as these locations account for the vast majority of burned area (Potter et al. 2023). Because of this our model testing data only includes fire polygons from 2001 to 2020 so that we could implement this active fire constraint. We considered active fire detections classified as nominal or high confidence, removing low confidence detections. More specifically we use MODIS active fire from 2001 to 2012, MODIS and VIIRS Suomi-NPP from 2012-present, MODIS/Suomi-NPP and VIIRS NOAA-20 from 2018 to 2023. It is common to buffer the hotspots to account for potential burned areas occurring outside these known hotspots, for example due to quickly spreading fires (Potter et al. 2023; Talucci, Loranty, and Alexander 2022; Veraverbeke, Rogers, and Randerson 2015). However, the influence of buffer size on estimated burned area needs to be considered. To assess this, we applied buffer zones ranging from 1 km to 8 km around active fire hits, allowing us to evaluate how varying sizes influenced model performance and estimates of total burned area.

Given that we constrain our model testing using active fire hits which are only available beginning in 2001, we needed to ensure only 2001–2020 polygons were used in our testing set used for final model performance evaluation. Due to this we used 80% of all polygons prior to 2001 in MTBS and NBAC as training and 20% as validation. We then used 70% of the 2001–2020 polygons from MTBS, NBAC and Talucci, Loranty, and Alexander (2022) as training, 10% as validation and 20% as testing (Table S4). We used a 5-fold cross-validation approach whereby a different 20% of the polygons from 2001 to 2020 were used in each of the test folds so that every single reference fire polygon was in the test set once, and our final error metrics were reported as the mean across these 5 folds. Since a different 20% of the 2001–2020 polygons is in the testing set in each of the 5 folds, we also shuffled the 70% and 10% of the polygons used in the 2001–2020 training and validation sets accordingly. The polygons used in the 80% of the data used to train/validate the model is never included in the 20% of the data used to test the model to ensure our test set and evaluation is completely independent. All error metrics reported from here on out represent the scores across the 5-fold cross-validation.

Lastly, we evaluated the performance of these models across the ABZ in addition to distinct ecoregions. In North America we used the Omernik and Griffith (2014) ecoregions, and in Eurasia we used Soja et al. (2004) ecoregions (Figure S1).

UNet++ models were trained for a maximum of 100 epochs, and parameters were tuned using Bayesian optimization (supplemental methods 1.2, Table S5). We used adaptive learning to reduce the learning rate if the validation IOU (Equation 5) did not increase in five consecutive iterations. If this condition was met, we reduced our learning rate by 50% for each subsequent iteration. We also implemented an early stopping procedure to prevent overfitting, whereby we stopped training and the model was

assumed to converge if its IOU did not improve for 20 consecutive iterations. IOU is calculated by dividing the intersection between the predicted and reference data by their union, which can also be expressed in terms of True Positives (TP), False Positives (FP) and False Negatives (FN).

$$\text{IOU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \tag{5}$$

An IOU score can range from zero to one, with one representing a perfect overlap between reference data and predictions. When IOU was maximized in the training/validation process and didn't improve for 20 consecutive iterations we then saved the weights of the model to be used on our held out test set and for use as final model evaluation. While IOU was the primary metric we used to determine if our model had stabilized in the training/validation process, when we evaluated our model on the test set we also calculated precision, recall, F1, and accuracy, which are functions of TP, FP, FN and True Negatives (TN). Precision measures the proportion of positive predictions made by the model that are correct:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{6}$$

Recall measures the proportion of actual positive cases that the model correctly identified:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{7}$$

F1 is the harmonic mean of precision and recall:

$$\text{F1} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{8}$$

Finally, accuracy is the total proportion of pixels predicted correctly:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{9}$$

We evaluated UNet++ models by comparing error metrics on the two image composite methods described in section 2.3 in three ways: 1) a combined North America and Eurasia model that was trained on all available training polygons, 2) a North American model trained on NBAC and MTBS, and 3) a Eurasia model trained on the Talucci, Loranty, and Alexander (2022) polygons. Because model performance was substantially weaker than the UNet++ (see Results), we only evaluated XGBoost models trained on the combined North American and Eurasian Dataset.

Lastly, since our UNet++ models output a probability that a pixel burned, the specific threshold chosen to classify a pixel as burned or unburned can substantially impact model performance. To identify the optimal threshold, we evaluated how precision and recall varied across a range of probability thresholds, assessing the trade-off between omission errors (false negatives) and commission errors (false positives). Higher thresholds typically yield greater precision by reducing false positives, whereas lower thresholds increase recall by capturing more true positives. Our analysis aimed to determine the threshold at which precision and recall are most effectively balanced.

2.4. Model inference

We implemented our best-performing model across the study domain while restricting our burned area to locations of MODIS/VIIRs active fire detections and also masked out croplands from any potential prediction location, as cropland management can display similar dNBR signals to fires and produce false positive detections. To identify croplands, we used the Potapov et al. (2022) global cropland layer, which is available for five epochs: 2001–2003, 2004–2007, 2008–2011, 2011–2015 and 2016–2019. We used the epoch most closely associated with a fire year of interest. We also removed water bodies using the Pekel et al. (2016) yearly water classification product.

Finally, we compared our burned area to ten other products ranging in resolution from 30 m to a quarter degree (supplemental methods 1.3, Table S6) for the years 2003, 2004, 2013, 2014, 2015 and 2023. We selected these years because they represent a range of large/small fire years across the domain. For example, 2003 is a very large fire year in Eurasia but small in North America, 2004 is large in North America but small in in Eurasia, and 2013, 2014 and 2015 are relatively large fire years in both. In addition, these years represent periods with varying satellite data availability: Landsat 5 and 7 were available only in 2003 and 2004; Landsat 7 and 8 in 2014; Landsat 7, 8, and Sentinel-2A in 2015; and Landsat 7, 8, 9, and Sentinel-2A and 2B in 2023.

3. Results

3.1. Influence of active fire buffer size on model performance

We evaluated error metrics on a range of active fire buffer sizes to determine how the choice of buffer size influenced model performance (Figure 3). For this comparison we used a multi-monthly composite UNet++ model trained on the combined North America and Eurasia reference fire polygons. We found precision decreased as buffer size increased, indicating that more false positives were present with larger buffer sizes. However, recall increased as the buffer size increased, indicating there were also fewer false negatives. Both the IOU and F1 scores, which incorporate the tradeoffs of precision and recall, plateau with a buffer size of approximately 6 km, after which the increase in overall performance is minimal. Therefore, for all future model comparisons we used a buffer size of 6 km.

The limitation of smaller buffer sizes is further illustrated by visually comparing predicted burned areas with different active fire buffer sizes to the reference polygons used for model training (Figures S5–S6). With smaller buffers, burned areas can be omitted due to spatial mismatches between actual burned locations and recorded active fire detections. Consequently, smaller buffer sizes fail to capture portions of fires that occurred beyond the immediate vicinity of active fire pixels.

3.2. XGboost vs. UNet++

We compared model performance on the 10% subset of reference polygons sampled from the 3000 randomly selected fires (Section 2.5), as well as across the annual and multi-monthly composite approaches for both the XGBoost and UNet++ models (Table 1). The UNet++ models performed considerably better than the XGBoost models regardless of

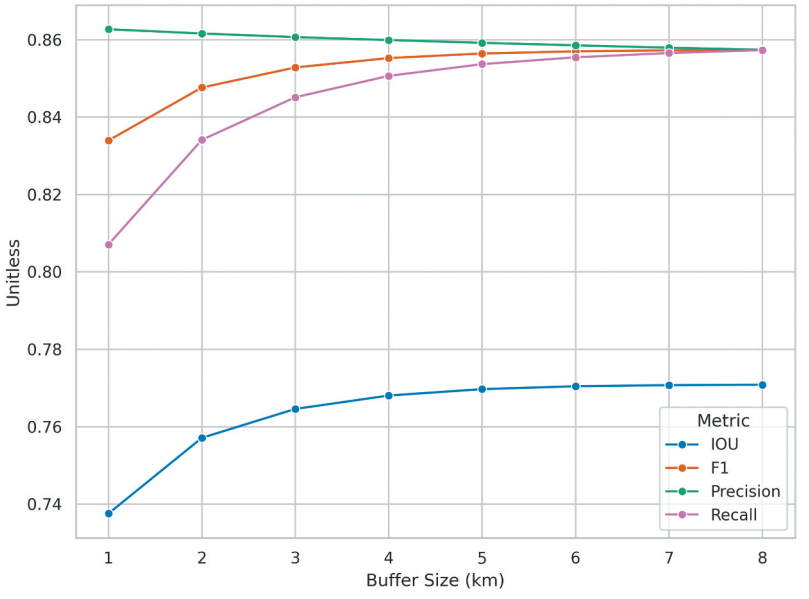

Figure 3. Comparison of intersection over union (IOU), F1, precision and recall across a range of active fire buffer sizes in the Arctic boreal zone (ABZ).

Table 1. Comparison of XGBoost and UNet++ models by image compositing method. Metrics used are IOU, precision, recall, F1, and accuracy on the testing set. Best scores are shown in bold.

Composite Method	Model	IOU	Precision	Recall	F1	Accuracy
Annual	XGBoost	0.64	0.88	0.71	0.78	0.94
Annual	UNet++	0.80	0.92	0.86	0.89	0.95
multi-monthly	XGBoost	0.65	0.88	0.72	0.79	0.95
multi-monthly	UNet++	**0.83**	**0.94**	**0.88**	**0.91**	**0.97**

image composite method (e.g. IOU of 0.80–0.83 vs. 0.64–0.65). Moreover, the multi-monthly method resulted in a slightly higher IOU than annual compositing for both XGBoost and UNet++.

3.3. UNet++ model performance by region

Because of their superior performance, we limited further evaluation and comparisons to the UNet++ models only. All error metrics were evaluated on the testing set of the 5-fold cross-validation. We then compared our regional UNet++ models trained on three regions: (1) combined North American and Eurasian fires, (2) North American only fires and (3) Eurasian only fires. Within each of these three regions we also compared our two image compositing methods (Table 2).

The multi-monthly composite method performed better than the annual composite method regardless of the training region. Regionally, the North American multi-monthly model performed the best (IOU = 0.83) while the Eurasian model performed somewhat worse (IOU = 0.71). The combined model (IOU = 0.77) performed worse than the North

Table 2. Comparison of UNet++ models by image composite methods and training regions. Metrics used are IOU, precision, recall, F1, and accuracy on the testing set. Best scores are shown in bold.

Composite Method	Region	IOU	Precision	Recall	F1	Accuracy
Annual	Combined	0.72	0.93	0.76	0.83	0.97
Annual	North America	0.76	0.90	0.83	0.87	0.97
Annual	Eurasia	0.68	0.89	0.77	0.83	0.96
multi-monthly	Combined	0.77	0.86	0.85	0.85	0.98
multi-monthly	North America	**0.83**	**0.93**	**0.88**	**0.91**	**0.97**
multi-monthly	Eurasia	0.71	0.89	0.78	0.83	0.96

Table 3. Comparison of the UNet++ model with multi-monthly image compositing by training and evaluation region. Metrics compared include IOU, precision, recall, F1, and accuracy on the testing set. Best scores are shown in bold.

Training Region	Evaluation Region	IOU	Precision	Recall	F1	Accuracy
Combined	North America	**0.84**	**0.89**	**0.94**	**0.91**	**0.98**
North America	North America	0.83	0.93	0.88	0.91	0.97
Combined	Eurasia	0.72	0.88	0.80	0.84	0.96
Eurasia	Eurasia	0.71	0.89	0.78	0.83	0.96

American model but better than the Eurasian model. Our models have high precision, meaning that when a fire is predicted, it's likely to be correct. However, their lower recall indicates that some fires are missed, suggesting the presence of false negatives. Regardless of the region, precision scores tend to be high, but recall scores are more variable, particularly in Eurasia which has lower recall than in North America. It is also worth noting that precision is somewhat lower when using the multi-monthly compositing method compared to the annual method for the combined Eurasia and North American models, but recall in the multi-monthly compositing method is considerably higher than the annual method. This indicates the multi-monthly compositing method's improved performance is due to a reduction in false negatives – i.e. omissions are reduced. A visual comparison of the annual and multi-monthly compositing methods demonstrates that the annual method can nearly miss entire fires (Figure S7 a,b,e,f). Additionally, even within fires identified by both methods, the multi-monthly compositing method more accurately captures the full extent of burned areas (Figure S7 i,j,m,n).

Table 2 highlights regional discrepancies in model performance, but it does not clarify whether a model trained on combined data from both regions outperforms models trained solely on regional data. In other words, it does not indicate whether a combined North America – Eurasia model performs better in North America than a North America-only model, or better in Eurasia than a Eurasia-only model. To investigate this, we compared performance between region-specific models and the combined model for each region separately (North America and Eurasia) using the multi-monthly compositing method (Table 3).

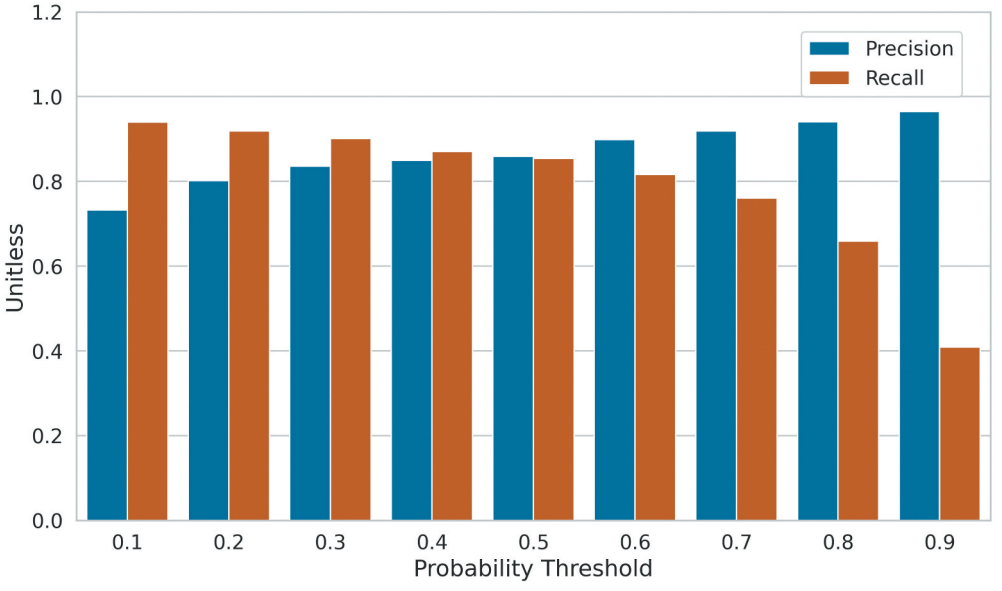

Figure 4. Comparison of precision and recall for the combined multi-monthly model within a 6 km active fire buffer across different probability thresholds and in the Arctic-boreal zone (ABZ).

Based on this comparison, a combined model does just as well in North America (IOU = 0.84) as a North American-only model (IOU = 0.83), and just as well in Eurasia (IOU = 0.72) as a Eurasian-only model (IOU = 0.71). Interestingly, precision scores for the combined model are nearly identical between North America and Eurasia, yet recall scores are different (0.94 for North America vs. 0.80 for Eurasia). Hence, the decreased model performance in Eurasia compared to North America is due almost exclusively to a larger number of false negatives, or omissions. As there is no benefit to using regional-based models, all remaining results focus on a combined multi-monthly model.

3.4. Effects of UNet++ probability threshold on model performance

To determine the optimal probability threshold for model implementation, we evaluated how precision and recall changed across various thresholds on the test set (Figure 4). For our model, precision and recall were most balanced at a probability threshold of 0.5. We therefore use a 0.5 probability threshold for model implementation.

3.5. UNet++ model performance by ecoregion

We also evaluated IOU, precision, recall and F1 scores across North American and Eurasian ecoregions to assess model performance at finer spatial scales. Using our combined multi-monthly image compositing model, it is apparent that IOU scores are variable depending on the ecoregion (Figure 5a), particularly within Eurasia, where the IOU scores are the highest in Central Siberia (0.79–0.84) and lower in ecoregions further to the south (0.52–0.59). IOU scores in North America are relatively consistent (0.84–0.90).

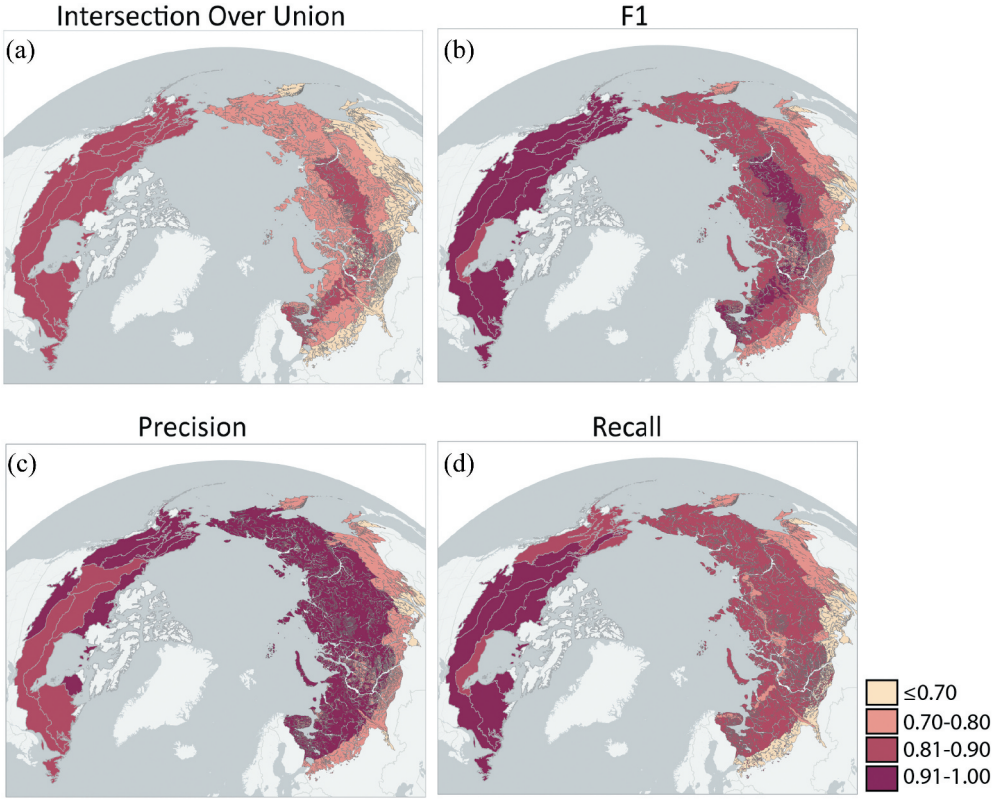

Figure 5. Comparison of the intersection over union (a), F1 (b), precision (c) and recall (d) for fires within individual ecoregions using the combined multi-monthly model.

Compared to the combined annual image compositing method, every ecoregion has a higher IOU score using the multi-monthly method (Figure 6a). Although the annual method has slightly higher precision in most ecoregions in both North America and Eurasia (Figure 6c), this difference is offset by the larger recall scores using the multi-monthly composite method (Figure 6d), and therefore stronger F1 scores overall (Figure 6b).

Given the variation in IOU scores between ecoregions, we compared the association between fire size and IOU for the combined model using both the annual and multi-monthly compositing methods for fires less than 120 ha in size (Figure S8). Overall, the association between fire size and IOU is weak ($R^2 < 0.01$). However, for larger fires within the training data ($\geq$ 25 ha), the IOU scores are consistently higher, particularly when using the multi-monthly image compositing method. Since the combined multi-monthly model performs better than the annual method, this was our chosen model for final implementation and all further analysis.

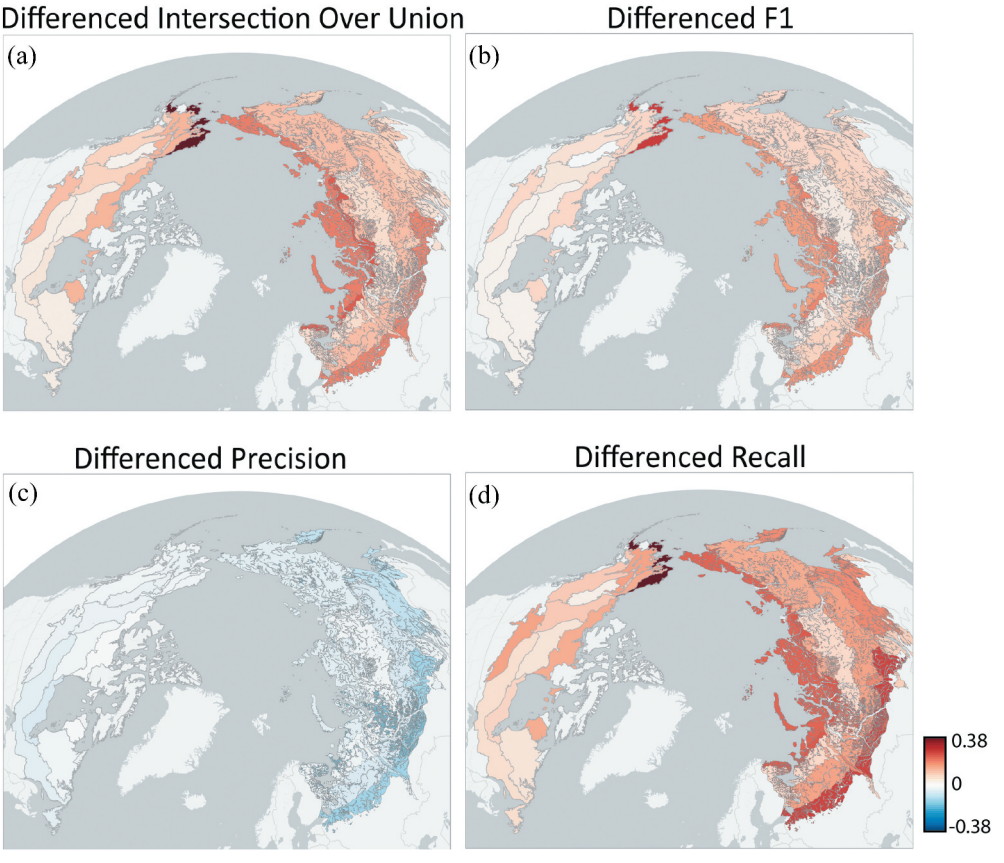

Figure 6. Comparison of the differenced intersection over union (a), F1 (b), precision (c) and recall (d) for fires within individual ecoregions. Differences are calculated by subtracting the combined annual compositing method model from the combined multi-monthly image compositing method model.

3.6. Burned area comparisons

We compared our predicted burned area to other products across the entire ABZ for the years 2003, 2004, 2013, 2014, 2015 and 2023 (Figure 7). Note that regional products not covering the entire domain were excluded here.

Although burned area is variable by product and year, on average GABAM predicts the most burned area and MCD64A1 the least. Our predictions are generally closest to ABBA and FireAtlas (Scholten et al. 2024). While not all products are available across all years, when the total burned area is compared for the years 2013–2015 (when all products have data; Figure S9), our predicted burned area is the median estimated burned area (34.26 Mha). The product with the most burned area (GABAM) predicts 71.2% more burned area (or 18.8 Mha) than the product with the least (MCD64A1).

We tend to predict slightly more burned area in North America than all other products besides GABAM and ABBA (Figures S10, S11). However, except for GABAM, which has considerably more burned area than all other products, the differences are relatively

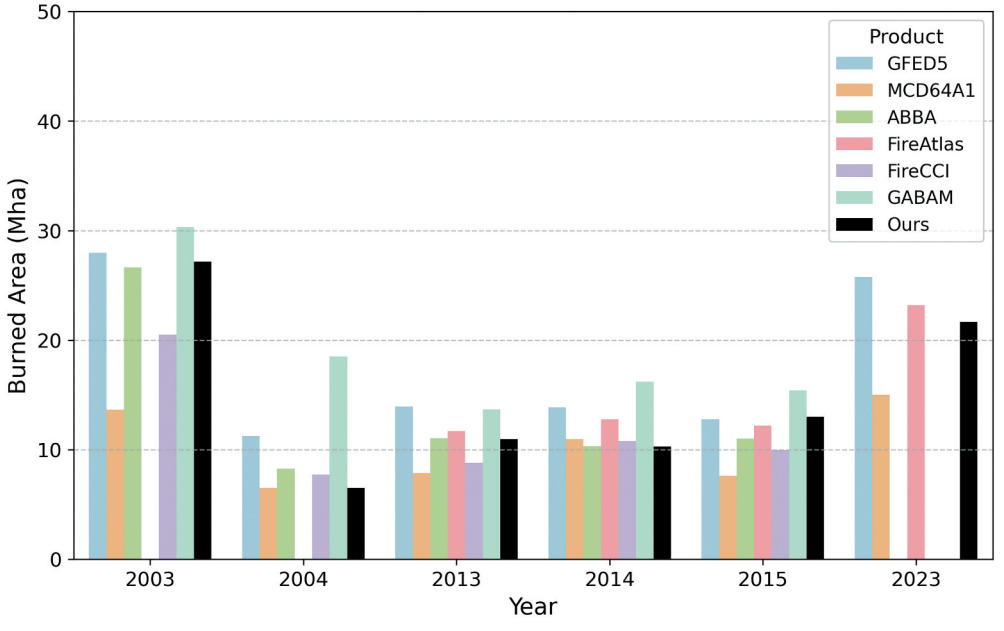

Figure 7. Comparison of predicted burned area between our model and other products (GFED5, MCD64A1, ABBA, FireAtlas, FireCCI and GABAM) across the study domain.

small. The results are similar to burned area comparisons in Canada (Figures S12, S13). Of particular note is the record-breaking fire year of 2023 in Canada, for which we predict 16.29 Mha of burned area; 1.73 Mha more than NBAC and 1.00 Mha less than FireAtlas. Our burned area predictions in Eurasia (Figures S14, S15) tend to be the median compared to other products, and closest in total burned area to FireAtlas. In Siberia (Figures S16, S17), we predict less burned area than Talucci, Loranty, and Alexander (2022),which includes unburned islands, and again closest in burned area estimates to FireAtlas and ABBA. In only Alaska (not shown), we predict 5.84% less burned area than MTBS, or 5.48 vs 5.82 Mha. Overall, our burned area estimates are within the range of previous products, although there is large inter-annual variability between products depending on year and region.

4. Discussion

Our final model, which used the combined multi-monthly compositing method, performed well on unseen test data, achieving an IOU score of 0.77 and an F1 score of 0.85 (Table 2). Notably, the multi-monthly compositing method outperformed the annual compositing method (IOU = 0.72), and the differences in approaches is visually distinct (Figure S7). It is likely the multi-monthly performs better than the annual compositing method because it is better at detecting understory fires as well as quickly recovering grassland or shrubland fires which burn early or late in the season, and therefore are already fully recovered one year postfire, which the annual method would not detect. This finding is particularly important given that previous burned area studies predominantly employed annual compositing techniques (Potter et al. 2023; Rogers et al. 2014; Talucci, Loranty, and Alexander 2022; Veraverbeke, Rogers, and Randerson 2015) or sliding

window approaches, such as that used in the MODIS MCD64A1 product, to detect maximum spectral change. To our knowledge, this study is the first to directly compare a multi-monthly compositing method against the widely used annual approach. Utilizing imagery from within the year of fire occurrence poses challenges, primarily due to difficulties in accurately distinguishing snow cover. This issue can introduce substantial noise into satellite-derived spectral signals, especially in northern regions. Here, we demonstrated that by implementing a snow-masking strategy based on the NDSI, it is possible to effectively remove snow-related interference, even in shoulder-season months. This enables accurate fire detection using multi-monthly compositing. Moreover, this multi-monthly compositing approach could potentially be adapted for use in other biomes that experience similar seasonal limitations due to snow cover.

Even with the multi-monthly compositing approach there were notable differences in model performance depending on the region of evaluation. Our model achieved an IOU of 0.84 in North America but only 0.72 in Eurasia (Table 3). The reasons for this are two-fold. First, the reference polygons provided by MTBS/NBAC for North America are likely more accurate than any available data in Eurasia, including the source we used (Talucci, Loranty, and Alexander 2022). More specifically, NBAC, which constitutes most of the reference training data in North America, accounts for unburned vegetation islands, while MTBS does not. In Eurasia, the variation in IOU by ecoregion is more pronounced than in North America (Figure 5a), likely because the fire regimes vary more in terms of seasonality, type of fire (crown vs surface), and dominant land cover types. In Eurasia, based on visual inspection of numerous reference polygons, the Talucci, Loranty, and Alexander (2022) data also often do not account for unburned vegetation patches. This is reflected by comparing the precision and recall scores in each region. In North America we have higher recall (0.94) than precision (0.89; Table 3), which shows we have slightly higher rates of false positives (commission error) than false negatives (omission error). In Eurasia we have higher precision (0.88) than recall (0.80; Table 3), which shows we have higher rates of false negatives than false positives. However, as there is some doubt as to how well the reference data initially accounts for unburned vegetation islands, these scores may not fully reflect reality. It is likely that the inclusion of unburned vegetation islands in the Eurasian testing data artificially deflates our accuracy scores, with the recall score in particular being lower than expected with a perfect training dataset. This is further confirmed by comparing our predicted burned area only in the region where Talucci, Loranty, and Alexander (2022) data exist (Figures 9, S16). We predict considerably less burned area than the reference dataset, and are much more in line with the alternative products. Secondly, the fire regime in North America is more amenable to satellite-based mapping than in Eurasia due to its higher frequency of stand-replacing fire compared to the lower-severity surface fires that dominate Eurasia (de Groot et al. 2013; Rogers et al. 2015). Although we were able to improve the model by using our multi-monthly image compositing method instead of the annual method (Figure 6), there is still room for improvement. One possible way to improve burned area estimates in this region would be to use the full time series of Landsat/Sentinel-2 imagery in a model that incorporates both space and time, such as a 3-dimensional CNN or a Long Short Term Memory neural network, fused with a CNN. This could be further facilitated by incorporating satellite-based radar information, which is not as sensitive to cloud cover as optical imagery and thus can provide a more complete time series. It is possible that even the multi-monthly

time steps used in this study are not sufficient to fully capture these fast recovering fires. It is possible that these methods may be particularly beneficial for mapping fires in southern Eurasia where we have comparatively lower recall scores (Figure 5) which indicates the presence of false negatives and an underestimation of burned area in this region.

We tend to predict much less burned area than GABAM (Figures 7, S9), which is the only other publicly available 30 m product that covers the ABZ. It has previously been documented that GABAM includes a high rate of false positives. In particular, there are occasions where entire Landsat scenes seem to be classified as burned, which increases the total burned area (Clelland et al. 2024). Overall, our burned area estimates across the ABZ are closest to FireAtlas and ABBA. While ABBA is a 500 m MODIS product, it was developed specifically for the Arctic-boreal region, as opposed to the global MCD64A1 product, which is likely why the burned area estimates in ABBA are closer to our own.

In Alaska and Canada we predict more burned area than the reference MTBS and NBAC datasets (Figures S10, S11). This is related to our higher recall than precision scores in North America, which indicate a higher commission error. In particular, by visually comparing to NBAC (Figure S18), we can see that there are occasions where we miss unburned islands that NBAC detects, but this phenomenon is variable by fire perimeter (Figures S19, S20). In addition, we also predict burned area within the active fire buffers that NBAC does not include. Across the six chosen years for wall-to-wall implementation of our model, 14.8% of our total burned area in Canada came from areas outside of the NBAC polygons but within the 6 km active fire buffers.

In Alaska, comparisons to the MTBS dataset (Figures 8, S21) demonstrate that our model more accurately incorporates unburned vegetation islands because unlike our approach, MTBS fire perimeters typically do not differentiate these unburned islands within their boundaries.

Southern Eurasia experienced particularly extensive wildfires in 2003, which was consistently observed across all products analysed (Figure S22). Nonetheless, our product

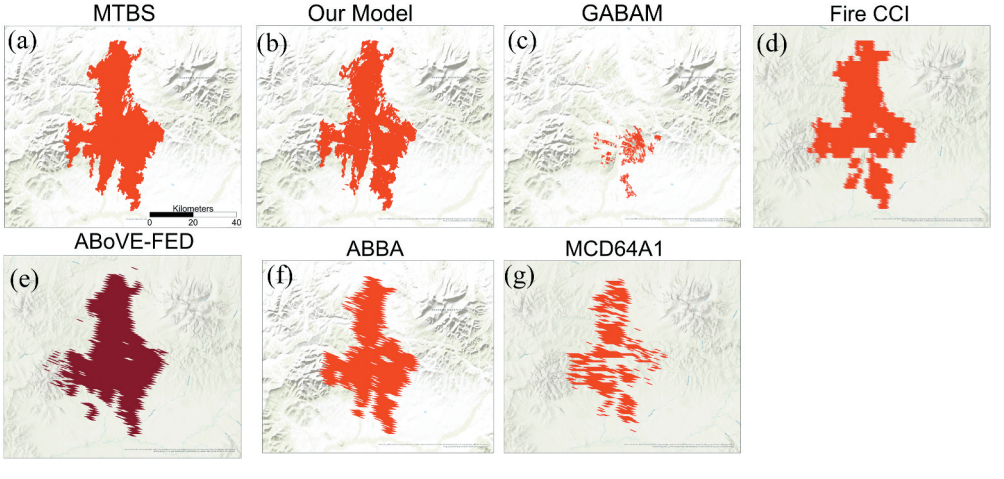

Figure 8. Comparison of burned area for an example 2005 fire in Alaska according to MTBS, our model, GABAM, fire CCI, ABoVE-FED, ABBA and MCD64A1. For ABoVE-FED, yellow colors are closer to 0 and red colors are closer to 100% burned fraction for each pixel. For all other products burned area is in red.

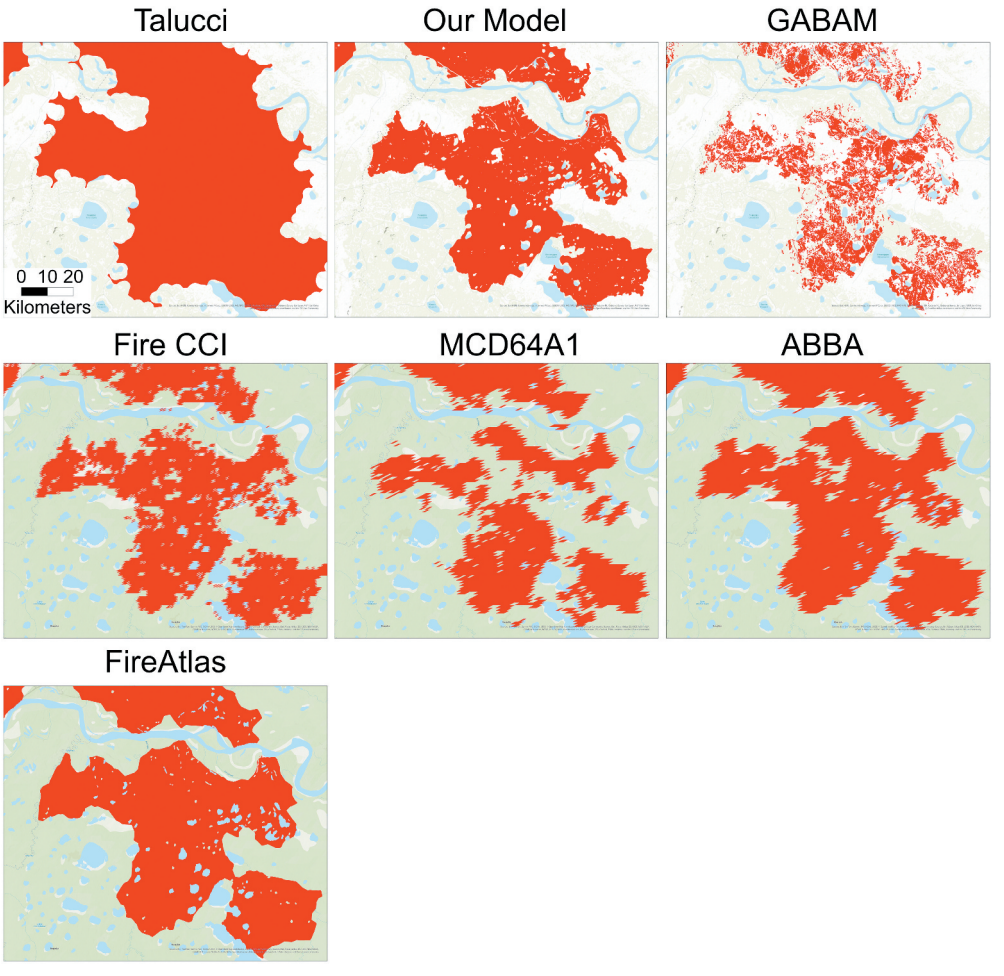

Figure 9. Comparison of burned area for a set of 2015 fires in Siberia according to Talucci et al., our model, GABAM, fire CCI, MCD64A1, ABBA and FireAtlas.

frequently estimates less burned area compared to the reference dataset (Talucci, Loranty, and Alexander 2022, Figure 9), which is due to our identification of unburned vegetation islands.

There are several difficulties that need to be considered when applying the approach we used. Firstly, while we used the most accurate reference training data available when training our model, neither MTBS nor Talucci, Loranty, and Alexander (2022) consistently accounted for unburned vegetation islands. Hence, our model is limited by the accuracy of the training data, and it is likely that we still include more unburned islands than truly exist. There are methods which will propagate the error of the reference datasets through to the final product (Foody 2023), yet these methods require that the accuracy of the reference datasets themselves are known. Of the three reference datasets we use the only accuracy metric reported is by NBAC, and we do compare our burned area to this product when appropriate. Despite this limitation, our model still captures unburned islands that

reference data sets do not (Figure 8, 9, S21). One approach that may further increase the accuracy of a 30 m product would be to use the raster version of MTBS, which includes an unburned/low severity fire class, and remove those locations from the MTBS polygons. While this approach would remove some areas that truly burned at low severity, it is possible the accuracy of classifying unburned islands would improve. Secondly, to limit false positives we constrained our predictions to be within 6 km of MODIS or VIIRS active fire hits. Without the use of this constraint, our burned area estimates would be considerably higher (Figure S23) and arguably less accurate. The use of the active fire hits and the selection of the buffer size itself influenced estimated burned area. If the buffer size is too small, burned area will be missed (Figures S5, S6); however, if the buffer size is too large, the rate of false positive detection will increase.

Another consideration in our approach is the temporal variability in satellite image availability, due to additional satellites becoming operational closer to the present. Specifically, from 2001 to 2012 only Landsat 5 and 7 imagery is available, from 2013 to 2016 imagery from Landsat 7 and 8 is available, and from 2017 to 2020 imagery from Landsat 7, 8, and Sentinel-2A and 2B is available. To assess whether these differences in satellite availability influenced our model results over time, we compared median IOU scores across these intervals (Table S7). However, we found no substantial temporal variability, with median IOU scores consistently ranging between 0.77 and 0.79.

Lastly, an inherent limitation of our multi-monthly image compositing approach is that we can't deploy our model, and therefore estimate burned area, until we have Landsat/Sentinel-2 imagery available one year post fire (for any given fire year) through August 31st (Figure S3). For applications which need near real time burned area estimates our method will not be useful and alternative burned area products will need to be used instead. As of this writing we can generate burned area from 2001 to 2024 because summer imagery from Landsat and Sentinel-2 has been aquired through August 31st of the year 2025. If, however, a pure multi-monthly approach was used whereby imagery 1-year post fire was not still used in the post-fire image compositing (Figure S3), it may be possible to use this approach to detect burned areas in near-real time as well.

Despite these considerations, our model performs well overall, producing burned area predictions that align with and, in most cases, improve upon previous estimates. Even though portions of our reference training data included unburned vegetation islands, our approach notably improves upon earlier methods by accurately capturing these unburned islands within the majority of the fire scars analysed. Furthermore, our use of high-resolution (30 m) imagery for model training results in significantly clearer and more precise fire perimeters compared to earlier estimates based on MODIS and VIIRS data. The ability to accurately map wildfires at 30 m resolution holds promise for numerous applications, including analysing the feedbacks between vegetation and burn patterns, assessing short-interval reburns, characterizing spatial patterns and dynamics of unburned islands, quantifying fine-scale impacts on wildlife habitats, evaluating post-fire erosion risks, guiding targeted rehabilitation and restoration efforts, and improving carbon emission estimates through detailed burn severity assessments.

5. Conclusion

In this study we developed a deep learning approach based on a UNet++ model to map wildfires at 30 m resolution across the ABZ. Evaluated on the years 2001–2020 our approach yielded promising results, achieving an overall IOU of 0.77. However, model performance varied notably by region, with better performance observed in North America (IOU = 0.84) compared to Eurasia (IOU = 0.72).

Our model's performance in North America is comparable to results reported by previous studies (Zhang et al. 2024). In contrast, this research represents the first application of such a method in Eurasia, marking a significant advancement in mapping burned areas across the entire ABZ. Overall, our burned area estimates closely align with MODIS-based estimates specifically tailored to the ABZ (Loboda et al. 2024), as well as with those derived from VIIRS active fire-based mapping (Scholten et al. 2024). However, when compared to the only other available 30 m resolution mapping approach (Long et al. 2019), our model predicted substantially lower burned area, primarily due to the higher rate of false positives that inflated burned area estimates in the previous approach (GABAM). Moreover, we better resolved areas within fire boundaries that escaped burning. While we only compared burned area across the ABZ for six selected fire years, future research steps include producing an annual burned area product from 2001 to 2023 for the ABZ that can be updated annually.

Acknowledgments

Computing resources for this work were provided by the NASA High-End Computing Program through the NASA Center for Climate Simulation at Goddard Space Flight Center.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was funded by the Gordon and Betty Moore Foundation (grant number 8414), with support from funding catalyzed by the Audacious Project (Permafrost Pathways), NSF Arctic System Science (award number 2116864) and Arctic Natural Sciences (award number 2019485), and the NASA Arctic-Boreal Vulnerability Experiment [80NSSC22K1244]. The contribution of S.V. was supported by the European Research Council through a Consolidator grant under the European Union's Horizon 2020 research and innovation program [grant number 101000987]. A.A.C. was funded by the IAPETUS2 Doctoral Training Partnership scheme through the UK Natural Environment Research Council [grant NE/S007431/1].

ORCID

Sander Veraverbeke http://orcid.org/0000-0003-1362-5125
Scott J. Goetz http://orcid.org/0000-0002-6326-4308
Michelle C. Mack http://orcid.org/0000-0003-1279-4242

Author contributions

S.P. and B.M.R. contributed to conceptualization. S.P., Y.Y., A.B., A.T., A.A.C. and B.M.R. contributed to methodology, investigation, and original draft writing. S.P. contributed formal analysis, data curation, and visualization. B.M.R. contributed to supervision and funding acquisition. S.N. contributed to funding acquisition. All authors contributed to the review and editing process.

Data availability

Any of the datasets presented in this manuscript are available upon request from the corresponding author. Code used in the creation of this manuscript is archived at the Arctic Data Center (Potter and Rogers 2025).

References

ACIA (Arctic Climate Impact Assessment). 2004. Impacts of a Warming Arctic: Arctic Climate Impact Assessment. Cambridge: Cambridge University Press. ACIA Overview report.

Amiro, B., J. Todd, B. Wotton, K. Logan, M. Flannigan, B. Stocks, J. Mason, D. Martell, and K. G. Hirsch. 2001. "Direct Carbon Emissions from Canadian Forest Fires, 1959-1999." *Canadian Journal of Forest Research* 31 (3): 512–525. https://doi.org/10.1139/x00-197.

Balshi, M. S., A. D. McGuire, P. Duffy, M. Flannigan, D. W. Kicklighter, and J. Melillo. 2009. "Vulnerability of Carbon Storage in North American Boreal Forests to Wildfires During the 21st Century." *Global Change Biology* 15 (6): 1491–1510. https://doi.org/10.1111/j.1365-2486.2009.01877.x.

Berner, L. T., J. J. Assmann, S. Normand, and S. J. Goetz. 2023. "'LandsatTS': An R Package to Facilitate Retrieval, Cleaning, Cross-Calibration, and Phenological Modeling of Landsat Time Series Data." *Ecography* 2023 (9): e06768. https://doi.org/10.1111/ecog.06768.

Berner, L. T., and S. J. Goetz. 2022. "Satellite Observations Document Trends Consistent with a Boreal Forest Biome Shift." *Global Change Biology* 28 (10): 3275–3292. https://doi.org/10.1111/gcb.16121.

Berner, L. T., R. Massey, P. Jantz, B. C. Forbes, M. Macias-Fauria, I. Myers-Smith, T. Kumpula, G. Gauthier, L. Andreu-Hayles, B. V. Gaglioti, et al. 2020. "Summer Warming Explains Widespread but Not Uniform Greening in the Arctic Tundra Biome." *Nature Communications* 11 (1): 4621. https://doi.org/10.1038/s41467-020-18479-5.

Bond-Lamberty, B., S. D. Peckham, D. E. Ahl, and S. T. Gower. 2007. "Fire as the Dominant Driver of Central Canadian Boreal Forest Carbon Balance." *Nature* 450 (7166): 89–92. https://doi.org/10.1038/nature06272.

Cao, X., J. Chen, B. Matsushita, H. Imura, and L. Wang. 2009. "An Automatic Method for Burn Scar Mapping Using Support Vector Machines." *International Journal of Remote Sensing* 30 (3): 577–594. https://doi.org/10.1080/01431160802220219.

Chen, T., and C. Guestrin. 2016. "Xgboost: A Scalable Tree Boosting System." In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794 doi:https://doi.org/10.1145/2939672.2939785. ACM.

Chen, Y., J. Hall, D. Van Wees, N. Andela, S. Hantson, L. Giglio, G. R. Van Der Werf, D. C. Morton, and J. T. Randerson. 2023. "Multi-Decadal Trends and Variability in Burned Area from the 5th Version of the Global Fire Emissions Database (GFED5)." *Earth System Science Data Discussions* 15:11 5227–5259.

Chuvieco, E., M. Pettinari, J. Lizundia-Loiola, T. Storm, and M. Padilla Parellada. 2018. *ESA Fire Climate Change Initiative (Fire_cci): MODIS Fire_cci Burned Area Pixel Product, Version 5.1* (CEDA Archive). November 1, 2018. https://doi.org/10.5285/58f00d8814064b79a0c49662ad3af537.

Chuvieco, E., C. Yue, A. Heil, F. Mouillot, I. Alonso-Canas, M. Padilla, J. M. Pereira, D. Oom, and K. Tansey. 2016. "A New Global Burned Area Product for Climate Assessment of Fire Impacts." *Global Ecology & Biogeography* 25 (5): 619–629. https://doi.org/10.1111/geb.12440.

Clelland, A. A., G. J. Marshall, R. Baxter, S. Potter, A. C. Talucci, J. M. Rady, H. Genet, B. M. Rogers, and S. M. Natali. 2024. "Annual and Seasonal Patterns of Burned Area Products in Arctic-Boreal North America and Russia for 2001–2020." *Remote Sensing* 16 (17): 3306. https://doi.org/10.3390/rs16173306.

de Groot, W. J., A. S. Cantin, M. D. Flannigan, A. J. Soja, L. M. Gowman, and A. Newbery. 2013. "A Comparison of Canadian and Russian Boreal Forest Fire Regimes." *Forest Ecology & Management* 294:23–34. https://doi.org/10.1016/j.foreco.2012.07.033.

Dinerstein, E., D. Olson, A. Joshi, C. Vynne, N. D. Burgess, E. Wikramanayake, N. Hahn, S. Palminteri, P. Hedao, R. Noss Hansen M. Locke H. Ellis E.C. Jones B. Barber C.V. Hayes R. Kormos C. Martin V. Crist E. Sechrest W. Price L., et al. 2017. "An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm." *Bioscience* 67 (6): 534–545. https://doi.org/10.1093/biosci/bix014.

Drusch, M., U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, et al. 2012. "Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services." *Remote Sensing of Environment* 120:25–36. https://doi.org/10.1016/j.rse.2011.11.026.

Eidenshink, J., B. Schwind, K. Brewer, Z.-L. Zhu, B. Quayle, and S. Howard. 2007. "A Project for Monitoring Trends in Burn Severity." *Fire Ecology* 3 (1): 3–21. https://doi.org/10.4996/fireecology.0301003.

Elmendorf, S. C., G. H. R. Henry, R. D. Hollister, R. G. Björk, A. D. Bjorkman, T. V. Callaghan, L. S. Collier, et al. 2012. "Global Assessment of Experimental Climate Warming on Tundra Vegetation: Heterogeneity Over Space and Time." *Ecology Letters* 15 (2): 164–175. https://doi.org/10.1111/j.1461-0248.2011.01716.x.

Fassnacht, F. E., E. Schmidt-Riese, T. Kattenborn, and J. Hernández. 2021. "Explaining Sentinel 2-Based DNBR and RDNBR Variability with Reference Data from the Bird's Eye (UAS) Perspective." *International Journal of Applied Earth Observation and Geoinformation* 95:102262. https://doi.org/10.1016/j.jag.2020.102262.

Flannigan, M., B. Stocks, M. Turetsky, and M. Wotton. 2009. "Impacts of Climate Change on Fire Activity and Fire Management in the Circumboreal Forest." *Global Change Biology* 15 (3): 549–560. https://doi.org/10.1111/j.1365-2486.2008.01660.x.

Foga, S., P. L. Scaramuzza, S. Guo, Z. Zhu, R. D. Dilley, T. Beckmann, G. L. Schmidt, J. L. Dwyer, M. J. Hughes, and B. Laue. 2017. "Cloud Detection Algorithm Comparison and Validation for Operational Landsat Data Products." *Remote Sensing of Environment* 194:379–390. https://doi.org/10.1016/j.rse.2017.03.026.

Foody, G. M. 2023. "Challenges in the Real World Use of Classification Accuracy Metrics: From Recall and Precision to the Matthews Correlation Coefficient. *PLOS ONE* 18 (10): e0291908. https://doi.org/10.1371/journal.pone.0291908.

French, N. H., L. K. Jenkins, T. V. Loboda, M. Flannigan, R. Jandt, L. L. Bourgeau-Chavez, and M. Whitley. 2015. "Fire in Arctic Tundra of Alaska: Past Fire Activity, Future Fire Potential, and Significance for Land Management and Ecology." *International Journal of Wildland Fire* 24 (8): 1045–1061. https://doi.org/10.1071/WF14167.

Furyaev, V. V., E. A. Vaganov, N. M. Tchebakova, and E. N. Valendik. 2001. "Effects of Fire and Climate on Successions and Structural Changes in the Siberian Boreal Forest." *Eurasian Journal of Forest Research* 2:1–15.

Gauthier, S., P. Bernier, T. Kuuluvainen, A. Z. Shvidenko, and D. G. Schepaschenko. 2015. "Boreal Forest Health and Global Change." *Science* 349 (6250): 819–822. https://doi.org/10.1126/science.aaa9092.

Genet, H., A. D. McGuire, K. Barrett, A. Breen, E. S. Euskirchen, J. Johnstone, E. S. Kasischke, et al. 2013. "Modeling the Effects of Fire Severity and Climate Warming on Active Layer Thickness and Soil Carbon Storage of Black Spruce Forests Across the Landscape in Interior Alaska." *Environmental Research Letters* 8 (4): 045016. https://doi.org/10.1088/1748-9326/8/4/045016.

Giglio, L. 2019. VIIRS/NPP Burned Area Monthly L4. Global 500m SIN Grid V001.

Giglio, L., L. Boschetti, D. P. Roy, M. L. Humber, and C. O. Justice. 2018. "The Collection 6 MODIS Burned Area Mapping Algorithm and Product." *Remote Sensing of Environment* 217:72–85. https://doi.org/10.1016/j.rse.2018.08.005.

Giglio, L., and C. Justice. 2015a. "MOD14A1 MODIS/Terra Thermal Anomalies/Fire Daily L3 Global 1km SIN Grid V006."

Giglio, L., and C. Justice. 2015b. "MYD14A1 MODIS/Aqua Thermal Anomalies/Fire Daily L3 Global 1km SIN Grid V006."

Giglio, L., J. T. Randerson, and G. R. Van Der Werf. 2013. "Analysis of Daily, Monthly, and Annual Burned Area Using the Fourth-Generation Global Fire Emissions Database (GFED4)." *Journal of Geophysical Research Biogeosciences* 118 (1): 317–328. https://doi.org/10.1002/jgrg.20042.

Goetz, S. J., M. C. Mack, K. R. Gurney, J. T. Randerson, and R. A. Houghton. 2007. "Ecosystem Responses to Recent Climate Change and Fire Disturbance at Northern High Latitudes: Observations and Model Results Contrasting Northern Eurasia and North America." *Environmental Research Letters* 2 (4): 045031. https://doi.org/10.1088/1748-9326/2/4/045031.

Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore. 2017. "Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone." *Remote Sensing of Environment* 202:18–27. https://doi.org/10.1016/j.rse.2017.06.031.

Guindon, L., P. Bernier, S. Gauthier, G. Stinson, P. Villemaire, and A. Beaudoin. 2018. "Missing Forest Cover Gains in Boreal Forests Explained." *Ecosphere* 9 (1): e02094. https://doi.org/10.1002/ecs2.2094.

Hall, R., R. Skakun, J. Metsaranta, R. Landry, R. Fraser, D. Raymond, M. Gartrell, V. Decker, and J. Little. 2020. "Generating Annual Estimates of Forest Fire Disturbance in Canada: The National Burned Area Composite." *International Journal of Wildland Fire* 29 (10): 878–891. https://doi.org/10.1071/WF19201.

Hardisky, M., R. Smart, and V. Klemas. 1983. "Seasonal Spectral Characteristics and Aboveground Biomass of the Tidal Marsh Plant." *Spartina-Alterniflora Photogrammetric Engineering and Remote Sensing* 49 (1): 85–92.

Hu, X., P. Zhang, and Y. Ban. 2023. "Large-Scale Burn Severity Mapping in Multispectral Imagery Using Deep Semantic Segmentation Models." *ISPRS Journal of Photogrammetry & Remote Sensing* 196:228–240. https://doi.org/10.1016/j.isprsjprs.2022.12.026.

Jain, P., S. C. Coogan, S. G. Subramanian, M. Crowley, S. Taylor, and M. D. Flannigan. 2020. "A Review of Machine Learning Applications in Wildfire Science and Management." *Environmental Reviews* 28 (4): 478–505. https://doi.org/10.1139/er-2020-0019.

Jia, G. J., I. C. Burke, A. F. Goetz, M. R. Kaufmann, and B. C. Kindel. 2006. "Assessing Spatial Patterns of Forest Fuel Using AVIRIS Data." *Remote Sensing of Environment* 102 (3–4): 318–327. https://doi.org/10.1016/j.rse.2006.02.025.

Johnson, E. A., and E. A. Johnson. 1996. *Fire and Vegetation Dynamics: Studies from the North American Boreal Forest*. Cambridge, UK: Cambridge University Press.

Johnstone, J. F., F. S. Chapin, T. N. Hollingsworth, M. C. Mack, V. Romanovsky, and M. Turetsky. 2010. "Fire, Climate Change, and Forest Resilience in Interior AlaskaThis Article is One of a Selection of Papers from the Dynamics of Change in Alaska's Boreal Forests: Resilience and Vulnerability in Response to Climate Warming." *Canadian Journal of Forest Research* 40 (7): 1302–1312. https://doi.org/10.1139/X10-061.

Jones, M. W., D. I. Kelley, C. A. Burton, F. Di Giuseppe, M. L. F. Barbosa, E. Brambleby, A. J. Hartley, et al. 2024. "State of Wildfires 2023–2024." *Earth System Science Data* 16 (8): 3601–3685. https://doi.org/10.5194/essd-16-3601-2024.

Kasischke, E. S., D. Williams, and D. Barry. 2002. "Analysis of the Patterns of Large Fires in the Boreal Forest Region of Alaska." *International Journal of Wildland Fire* 11 (2): 131–144. https://doi.org/10.1071/WF02023.

Kelly, R., M. L. Chipman, P. E. Higuera, I. Stefanova, L. B. Brubaker, and F. S. Hu. 2013. "Recent Burning of Boreal Forests Exceeds Fire Regime Limits of the Past 10,000 Years." *Proceedings of the National Academy of Sciences* 110 (32): 13055–13060. https://doi.org/10.1073/pnas.1305069110.

Key, C., and N. Benson. 2006. "La-1 Landscape Assessment (LA) Sampling and Analysis Methods." Firemon: Fire Effects Monitoring and Inventory System; US Department of Agriculture: Fort Collins, CO, USA.

Knopp, L., M. Wieland, M. Rättich, and S. Martinis. 2020. "A Deep Learning Approach for Burned Area Segmentation with Sentinel-2 Data." *Remote Sensing* 12 (15): 2422. https://doi.org/10.3390/rs12152422.

LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner. 1998. "Gradient-Based Learning Applied to Document Recognition." *Proceedings of the IEEE* 86 (11): 2278–2324. https://doi.org/10.1109/5.726791.

Li, F., D. M. Lawrence, and B. Bond-Lamberty. 2017. "Impact of Fire on Global Land Surface Air Temperature and Energy Budget for the 20th Century Due to Changes within Ecosystems." *Environmental Research Letters* 12 (4): 044014. https://doi.org/10.1088/1748-9326/aa6685.

Loboda, T., J. Hall, D. Chen, A. Hoffman-Hall, V. Shevade, F. Argueta, and X. Liang. 2024. "Arctic Boreal Annual Burned Area, Circumpolar Boreal Forest and Tundra." *Ornl Daac* 2:2002–2022.

Loboda, T., K. O'neal, and I. Csiszar. 2007. "Regionally Adaptable DNBR-Based Algorithm for Burned Area Mapping from MODIS Data." *Remote Sensing of Environment* 109 (4): 429–442. https://doi.org/10.1016/j.rse.2007.01.017.

Long, T., Z. Zhang, G. He, W. Jiao, C. Tang, B. Wu, X. Zhang, G. Wang, and R. Yin. 2019. "30 m Resolution Global Annual Burned Area Mapping Based on Landsat Images and Google Earth Engine." *Remote Sensing* 11 (5): 489. https://doi.org/10.3390/rs11050489.

McGuire, A. D., L. G. Anderson, T. R. Christensen, S. Dallimore, L. Guo, D. J. Hayes, M. Heimann, T. D. Lorenson, R. W. Macdonald, and N. Roulet. 2009. "Sensitivity of the Carbon Cycle in the Arctic to Climate Change." *Ecological Monographs* 79 (4): 523–555. https://doi.org/10.1890/08-2025.1.

McRae, D., S. Conard, G. Ivanova, A. Sukhinin, S. Baker, Y. Samsonov, T. Blake, et al. 2006. "Variability of Fire Behavior, Fire Effects, and Emissions in Scotch Pine Forests of Central Siberia." *Mitigation and Adaptation Strategies for Global Change* 11 (1): 45–74. https://doi.org/10.1007/s11027-006-1008-4.

Mekonnen, Z. A., W. J. Riley, J. T. Randerson, R. F. Grant, and B. M. Rogers. 2019. "Expansion of High-Latitude Deciduous Forests Driven by Interactions Between Climate Warming and Fire." *Nature Plants* 5 (9): 952–958. https://doi.org/10.1038/s41477-019-0495-8.

Miller, J. D., and A. E. Thode. 2007. "Quantifying Burn Severity in a Heterogeneous Landscape with a Relative Version of the Delta Normalized Burn Ratio (dNBR)." *Remote Sensing of Environment* 109 (1): 66–80. https://doi.org/10.1016/j.rse.2006.12.006.

Omernik, J. M., and G. E. Griffith. 2014. "Ecoregions of the Conterminous United States: Evolution of a Hierarchical Spatial Framework." *Environmental Management* 54 (6): 1249–1266. https://doi.org/10.1007/s00267-014-0364-1.

Payette, S. 1992. "Fire as a Controlling Process in the North American Boreal Forest." In *A System Analysis of the Global Boreal Forest* (Cambridge University Press), edited by HH. Shugart, R. Leemans, and GB. Bonan, 144–169.

Pekel, J.-F., A. Cottam, N. Gorelick, and A. S. Belward. 2016. "High-Resolution Mapping of Global Surface Water and Its Long-Term Changes." *Nature* 540 (7633): 418–422. https://doi.org/10.1038/nature20584.

Pereira, A. A., J. M. Pereira, R. Libonati, D. Oom, A. W. Setzer, F. Morelli, F. Machado-Silva, and L. M. T. De Carvalho. 2017. "Burned Area Mapping in the Brazilian Savanna Using a One-Class Support Vector Machine Trained by Active Fires." *Remote Sensing* 9 (11): 1161. https://doi.org/10.3390/rs9111161.

Phillips, C. A., B. M. Rogers, M. Elder, S. Cooperdock, M. Moubarak, J. T. Randerson, and P. C. Frumhoff. 2022. "Escalating Carbon Emissions from North American Boreal Forest Wildfires and the Climate Mitigation Potential of Fire Management." *Science Advances* 8 (17): eabl7161. https://doi.org/10.1126/sciadv.abl7161.

Potapov, P., S. Turubanova, M. C. Hansen, A. Tyukavina, V. Zalles, A. Khan, X.-P. Song, A. Pickens, Q. Shen, and J. Cortez. 2022. "Global Maps of Cropland Extent and Change Show Accelerated Cropland Expansion in the Twenty-first Century." *Nature Food* 3 (1): 19–28. https://doi.org/10.1038/s43016-021-00429-z.

Potter, S., S. Cooperdock, S. Veraverbeke, X. Walker, M. C. Mack, S. J. Goetz, J. Baltzer, et al. 2023. "Burned Area and Carbon Emissions Across Northwestern Boreal North America from 2001–2019." *Biogeosciences* 20 (13): 2785–2804. https://doi.org/10.5194/bg-20-2785-2023.

Potter, S., and B. Rogers. 2025. "Burned Area Mapping Across the Arctic-Boreal Zone (1985–2020) with Landsat and Sentinel-2 Imagery Arctic Data Center doi:10.18739/A2GB1XK04."

Ramo, R., and E. Chuvieco. 2017. "Developing a Random Forest Algorithm for MODIS Global Burned Area Classification." *Remote Sensing* 9 (11): 1193. https://doi.org/10.3390/rs9111193.

Randerson, J., Y. Chen, G. Van Der Werf, B. Rogers, and D. Morton. 2012. "Global Burned Area and Biomass Burning Emissions from Small Fires." *Journal of Geophysical Research Biogeosciences* 117 (G4). https://doi.org/10.1029/2012JG002128.

Rantanen, M., A. Y. Karpechko, A. Lipponen, K. Nordling, O. Hyvärinen, K. Ruosteenoja, T. Vihma, and A. Laaksonen. 2022. "The Arctic Has Warmed Nearly Four Times Faster Than the Globe Since 1979." *Communications Earth & Environment* 3 (1): 168. https://doi.org/10.1038/s43247-022-00498-3.

Riggs, G. A., D. K. Hall, and M. O. Román. 2015. "VIIRS Snow Cover Algorithm Theoretical Basis Document." Technical Report.

Rogers, B. M., A. J. Soja, M. L. Goulden, and J. T. Randerson. 2015. "Influence of Tree Species on Continental Differences in Boreal Fires and Climate Feedbacks." *Nature Geoscience* 8 (3): 228–234. https://doi.org/10.1038/ngeo2352.

Rogers, B. M., S. Veraverbeke, G. Azzari, C. Czimczik, S. Holden, G. Mouteva, F. Sedano, K. Treseder, and J. Randerson. 2014. "Quantifying Fire-Wide Carbon Emissions in Interior Alaska Using Field Measurements and Landsat Imagery." *Journal of Geophysical Research Biogeosciences* 119 (8): 1608–1629. https://doi.org/10.1002/2014JG002657.

Romanov, A. A., A. N. Tamarovskaya, E. Gloor, R. Brienen, B. A. Gusev, E. V. Leonenko, A. S. Vasiliev, and E. E. Krikunov. 2022. "Reassessment of Carbon Emissions from Fires and a New Estimate of Net Carbon Uptake in Russian Forests in 2001–2021." *Science of the Total Environment* 846:157322. https://doi.org/10.1016/j.scitotenv.2022.157322.

Roteta, E., A. Bastarrika, M. Padilla, T. Storm, and E. Chuvieco. 2019. "Development of a Sentinel-2 Burned Area Algorithm: Generation of a Small Fire Database for Sub-Saharan Africa." *Remote Sensing of Environment* 222:1–17. https://doi.org/10.1016/j.rse.2018.12.011.

Roy, D. P., H. Huang, L. Boschetti, L. Giglio, L. Yan, H. H. Zhang, and Z. Li. 2019. "Landsat-8 and Sentinel-2 Burned Area Mapping-A Combined Sensor Multi-Temporal Change Detection Approach." *Remote Sensing of Environment* 231:111254. https://doi.org/10.1016/j.rse.2019.111254.

Roy, D. P., V. Kovalskyy, H. Zhang, E. F. Vermote, L. Yan, S. Kumar, and A. Egorov. 2016. "Characterization of Landsat-7 to Landsat-8 Reflective Wavelength and Normalized Difference Vegetation Index Continuity." *Remote Sensing of Environment* 185:57–70. https://doi.org/10.1016/j.rse.2015.12.024.

Roy, D. P., M. A. Wulder, T. R. Loveland, R. G. Allen, M. C. Anderson, D. Helder, J. R. Irons, D. M. Johnson, R. Kennedy, T. A. Scambos, et al. 2014. "Landsat-8: Science and Product Vision for Terrestrial Global Change Research." *Remote Sensing of Environment* 145:154–172. https://doi.org/10.1016/j.rse.2014.02.001.

Scherer-Lorenzen, M., and E.-D. Schulze. 2005. *Forest Diversity and Function: Temperate and Boreal Systems*. Vol. 176. Berlin, DE: Springer Science & Business Media.

Scholten, R. C., S. Veraverbeke, Y. Chen, and J. T. Randerson. 2024. "Spatial Variability in Arctic–Boreal Fire Regimes Influenced by Environmental and Human Factors." *Nature Geoscience* 17 : 1–8.

Schroeder, W., P. Oliva, L. Giglio, and I. A. Csiszar. 2014. "The New VIIRS 375 m Active Fire Detection Data Product: Algorithm Description and Initial Assessment." *Remote Sensing of Environment* 143:85–96. https://doi.org/10.1016/j.rse.2013.12.008.

Serreze, M. C., and R. G. Barry. 2011. "Processes and Impacts of Arctic Amplification: A Research Synthesis." *Global & Planetary Change* 77 (1–2): 85–96. https://doi.org/10.1016/j.gloplacha.2011.03.004.

Shvidenko, A., and D. Schepaschenko. 2013. "Climate Change and Wildfires in Russia." *Contemporary Problems of Ecology* 6 (7): 683–692. https://doi.org/10.1134/S199542551307010X.

Shvidenko, A., D. Shchepashchenko, A. Sukhinin, I. McCallum, and S. Maksyutov. 2011. "Carbon Emissions from Forest Fires in Boreal Eurasia Between 1998-2010." In *Proceedings of the 5th International Wildland Fire Conference*, May (Sun City, South Africa: International Wildland Fire Conference), 9–13.

Skakun, S., J. Wevers, C. Brockmann, G. Doxani, M. Aleksandrov, M. Batič, D. Frantz, et al. 2022. "Cloud Mask Intercomparison Exercise (CMIX): An Evaluation of Cloud Masking Algorithms for Landsat 8

and Sentinel-2." *Remote Sensing of Environment* 274:112990. https://doi.org/10.1016/j.rse.2022.112990.

Soja, A. J., W. R. Cofer, H. H. Shugart, A. I. Sukhinin, P. W. Stackhouse, D. J. McRae, and S. G. Conard. 2004. "Estimating Fire Emissions and Disparities in Boreal Siberia (1998–2002)." *Journal of Geophysical Research Atmospheres* 109 (D14). https://doi.org/10.1029/2004JD004570.

Stocks, B., J. Mason, J. Todd, E. Bosch, B. Wotton, B. Amiro, M. Flannigan, K. Hirsch, K. Logan, D. Martell, et al. 2002. "Large Forest Fires in Canada, 1959–1997." *Journal of Geophysical Research Atmospheres* 107 (D1): FFR–5. https://doi.org/10.1029/2001JD000484.

Sulla-Menashe, D., M. A. Friedl, and C. E. Woodcock. 2016. "Sources of Bias and Variability in Long-Term Landsat Time Series Over Canadian Boreal Forests." *Remote Sensing of Environment* 177:206–219. https://doi.org/10.1016/j.rse.2016.02.041.

Talucci, A. C., M. M. Loranty, and H. D. Alexander. 2022. "Siberian Taiga and Tundra Fire Regimes from 2001–2020." *Environmental Research Letters* 17 (2): 025001. https://doi.org/10.1088/1748-9326/ac3f07.

Todd, J., K. Melanie, D. Joshua, L. Gail, T. Jeff, M. Nicole, K. Megan, J. Joshua, M. Stephen, L. John, et al. 2017. "Mapping Burned Areas Using Dense Time-Series of Landsat Data." *Remote Sensing of Environment* 198 (1): 504–522.

Tucker, C. J. 1979. "Red and Photographic Infrared Linear Combinations for Monitoring Vegetation." *Remote Sensing of Environment* 8 (2): 127–150. https://doi.org/10.1016/0034-4257(79)90013-0.

Van der Werf, G. R., J. T. Randerson, L. Giglio, G. Collatz, M. Mu, P. S. Kasibhatla, D. C. Morton, R. DeFries, Y. V. Jin, and T. T. van Leeuwen. 2010. "Global Fire Emissions and the Contribution of Deforestation, Savanna, Forest, Agricultural, and Peat Fires (1997–2009)." *Atmospheric Chemistry & Physics* 10 (23): 11707–11735. https://doi.org/10.5194/acp-10-11707-2010.

Veraverbeke, S., B. Rogers, and J. Randerson. 2015. "Daily Burned Area and Carbon Emissions from Boreal Fires in Alaska." *Biogeosciences* 12 (11): 3579–3601. https://doi.org/10.5194/bg-12-3579-2015.

Viereck, L. A. 1973. "Wildfire in the Taiga of Alaska." *Quaternary Research* 3 (3): 465–495. https://doi.org/10.1016/0033-5894(73)90009-4.

Virkkala, A.-M., B. M. Rogers, J. D. Watts, K. A. Arndt, S. Potter, I. Wargowsky, E. A. G. Schuur, et al. 2025. "Wildfires Offset the Increasing but Spatially Heterogeneous Arctic–Boreal CO2 Uptake." *Nature Climate Change* 15 (2): 188–195. https://doi.org/10.1038/s41558-024-02234-5.

Walker, X. J., J. L. Baltzer, S. G. Cumming, N. J. Day, C. Ebert, S. Goetz, J. F. Johnstone, et al. 2019. "Increasing Wildfires Threaten Historic Carbon Sink of Boreal Forest Soils." *Nature* 572 (7770): 520–523. https://doi.org/10.1038/s41586-019-1474-y.

Walker, X. J., B. M. Rogers, J. L. Baltzer, S. G. Cumming, N. J. Day, S. J. Goetz, J. F. Johnstone, E. A. G. Schuur, M. R. Turetsky, and M. C. Mack. 2018. "Cross-Scale Controls on Carbon Emissions from Boreal Forest Megafires." *Global Change Biology* 24 (9): 4251–4265. https://doi.org/10.1111/gcb.14287.

Wang, J. A., A. Baccini, M. Farina, J. T. Randerson, and M. A. Friedl. 2021. "Disturbance Suppresses the Aboveground Carbon Sink in North American Boreal Forests." *Nature Climate Change* 11 (5): 435–441. https://doi.org/10.1038/s41558-021-01027-4.

Young, A. M., P. E. Higuera, P. A. Duffy, and F. S. Hu. 2017. "Climatic Thresholds Shape Northern High-Latitude Fire Regimes and Imply Vulnerability to Future Climate Change." *Ecography* 40 (5): 606–617. https://doi.org/10.1111/ecog.02205.

Zhang, P., X. Hu, Y. Ban, A. Nascetti, and M. Gong. 2024. "Assessing Sentinel-2, Sentinel-1, and ALOS-2 PALSAR-2 Data for Large-Scale Wildfire-Burned Area Mapping: Insights from the 2017–2019 Canada Wildfires." *Remote Sensing* 16 (3): 556. https://doi.org/10.3390/rs16030556.

Zhou, Z., M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang. 2019. "Unet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation." *IEEE Transactions on Medical Imaging* 39 (6): 1856–1867. https://doi.org/10.1109/TMI.2019.2959609.