Prediction of Fire Severity in the Peatlands

Shreya Bali[†]
Carnegie Mellon University

Blair Chen[†]
Carnegie Mellon University

Anirban Chowdhury[†]
Carnegie Mellon University

Akshina Gupta[†]
Carnegie Mellon University

Yue Wu[†]
Carnegie Mellon University

Sydney Zheng[†]
Carnegie Mellon University

Abstract

Peat fires, the largest fires on earth in terms of fuel consumption, are responsible for a significant portion of global carbon emissions. Predicting fire in the peatlands could help decision makers and researchers prevent or monitor peat fires. Despite this, research on fire prediction for peatlands remains largely understudied as compared to other forms of fires. However, peatland fires are unique and therefore require datasets and architectures attuned to their particular characteristics. In this paper, we propose several models to tackle the problem of fire prediction for the peatlands. We adapt U-Net and UNet-LSTM neural network architectures for peat fire prediction. Furthermore, we develop novel neural architectures for peatland fire prediction, PeatNet and PT-Net, based upon a graph-based and an transformer-based approach, respectively. In addition, We also present a new dataset, PeatSet, designed specifically for the problem of peatland fire prediction using previously existing datasets in the region of Canada. Our results indicate that these new deep-learning based architectures outperform a regression baseline from existing peatland research. Among all the tested models, PT-Net achieves the highest F1 score and an overall accuracy of 99.84%.

1 Introduction

Peatlands are a type of wetland that include marshes, bogs, fens, and swamps. They sequester more than twice as much carbon as stored in the world's forests despite covering only 3% of the Earth's land area (International Union for Conservation of Nature, 2017; Turetsky et al., 2015). Peatlands comprise of peat, a carbon-rich organic soil made of plant material that has decomposed and accumulated over millennia in the wet, anaerobic, acidic, and nutrient-deficient ground conditions. Over time, layered peat soil grows thicker, sequestering more carbon (Turetsky et al., 2002).

Increasingly, peatland ecosystems are drying, giving way to disastrous peat fires. Under normal conditions, the naturally wet ecosystem preserves the peat soil and prevents fires. However, the peatland water table is lowering due both to human activity through drainage and to natural droughts, weakening the natural protections provided by moisture. Climate change has exacerbated the magnitude and frequency of fires and the length of the fire season (Flannigan et al., 2009), forming a positive feedback loop of peatland carbon emissions.

Peat fires release a large amount of the carbon sequestered in peatlands, emitting massive amounts of carbon dioxide. In 2015, the daily CO_2 emissions of Indonesian fires originating mostly from peatlands exceeded the daily emissions of the entire United States (Harris et al., 2015). Human drainage and burning of peatlands releases 1.3 gigatons of CO_2 per year, almost 6% of global anthropogenic CO_2 emissions annually. (International Union for Conservation of Nature, 2017).

[†]equal contribution

While peatlands exist in many climates, the largest portion of peatlands is found in the northern countries of Canada and Russia (Wieder and Vitt, 2006). Canada contains 27% of the world's peatlands and is more densely covered by peatlands than Russia (Short, Short). Furthermore, there are several fire resources made publicly available as a part of Canadian conservation efforts. By studying boreal peatlands, we can tackle problems that are particularly unique to peatland fires, such as detecting fires under snow. As such, we focus our data collection efforts on Canada. However, similar to other areas, there is no single dataset that correlates all these different data sources with specific emphasis on peatlands.

Several features of peatland fires differentiate them from commonly-explored forest fire approaches. First, peat fires can smoulder, burn underground, and produce less heat than typical fires, rendering them difficult to detect. Peat fires can occur in wet or cold areas, even under snow as exemplified by western Canadian peat fires (Thompson, 2020). Such behavior is contradictory to the simplifying assumptions adopted for fire detection. For example, the Fire Emissions from NCAR (FINN) dataset (Wiedinmyer et al., 2011) assumes that emissions in snow cannot be caused by fires, an assumption which fails to cover Canadian peat fires in particular. Second, peat fires can last for months when they are smouldering underground, challenging the use of burn duration as a heuristic for fire severity. Given the environmental impact of peat fires, CO_2 emissions should ideally serve as the best quantification of fire severity. Third, factors such as soil carbon and soil moisture are far more important for peatlands than accounted for by forest fire models since the soil itself is a fuel source. In general, understanding this unique biome requires an entirely different interpretation of environmental factors than for typical forests. Fourth, the spread pattern of peatland fires is characterized by their ability to go underground and resurface later. This phenomenon challenges the common fire behavior assumption that fire can only spread to adjacent areas. These properties cause normal fire prediction techniques to be less effective in the case of peat fires since their unique behavior and impact is not captured by most existing fire prediction work.

The unique nature of peatland fires therefore presents a valuable opportunity to understand a major source of carbon emissions. Prior works on wildfire modeling have focused mostly on forest fires (Jain et al., 2020), with little attention to peatland fires. Most prior works in the domain of peatland research have been limited in either performance or scalability (Dadap et al., 2019; Bourgeau-Chavez et al., 2020a,b; Kalacska et al., 2018; Widyatmanti et al., 2019; Honma et al., 2016; Listyorini and Rahim, 2018; Hugelius et al., 2020). For example, methods using simple statistical methods achieve limited accuracy, while higher-accuracy mathematical models or physics-based solutions are either too slow to be used real-time or require super-computing resources (Jain et al., 2020).

On the other hand, machine learning offers a solution to achieve both improved accuracy and efficiency. Simple linear machine learning solutions have shown promise with significantly fewer computational resources (Sitanggang et al., 2014; Maulana et al., 2019). Since the general relationship of predictors to fire occurrence and severity is potentially non-linear, we propose using non-linear, spatially and temporally aware models to predict burned area for a subsequent time step based upon the current environment. This functions as a fire spread model that simulates daily spread. We seek to output two-dimensional map-like predictions from our model. This would allow for detailed spatial information which could theoretically be used by governmental agencies or other interested parties to more effectively monitor peatlands.

Our first general contribution are novel neural network architectures that we propose, implement, and test. With these models we seek to improve upon current peat fire prediction methods and potentially regular forest fire prediction. First, we adapt U-Net (Ronneberger et al., 2015) for use in peat fire prediction to capture spatial information about fire spread. Second, we develop an architecture that integrates U-Net and long short-term memory (LSTM) layers in order to capture both the spatial and temporal aspects of fire behavior. Third, we create PeatNet, which combines recurrent and convolutional layers with graph neural layers. The use of graph neural layers in PeatNet helps capture amorphous spatial shapes and allows for custom connections between spatial areas that can better represent underground spread. Finally, we create PT-Net by combining the attentional transformers model (Vaswani et al., 2017) with a ResNet architecture (He et al., 2015) because attentional models are thought to deal with time-series data better than traditional LSTMs (Vaswani et al., 2017).

Our second contribution is the curation of PeatSet: a Canada-specific peatland dataset built from a collection of relevant datasets for peatland fire prediction. We use PeatSet to train and test our models. These datasets were selected after consultation from several experts working on peatlands, as well as directly contacting fire dataset creators. We hope our efforts in gathering datasets will benefit future studies regarding peatlands and highlight the increased need to monitor this precious land resource.

To the best of our knowledge, this is the first paper to have an in-depth discussion about the use of different deep learning methods to predict peat fires, and also the only paper that has used Graph Neural Networks or Transformer architectures for the task of wildfire behavior prediction in general.

2 Related Work

2.1 Forest and Urban Fire Prediction

Fire behavior prediction methods generally focus on predicting growth and spread or predicting final severity. In order to predict severity, many models use metrics such as final burned area or duration of burn. Using the former assumes severity increases when more land is burned, and using the latter assume severity increases the longer the fire lasts.

Common fire behavior prediction methods are regression (Castelli et al., 2015; Cortez and Morais, 2007; Mitsopoulos and Mallinis, 2017), random forests (Markuzon and Kolitz, 2009; Mitsopoulos and Mallinis, 2017), support vector machines (Castelli et al., 2015; Cortez and Morais, 2007), or Bayesian networks (Markuzon and Kolitz, 2009). We use both regression as a baseline.

Researchers have traditionally formulated the fire prediction problem as a classification problem. The highest accuracy among these models is 97.5% and is achieved by Shidik and Mustofa (2014). Similar research achieves far lower accuracy (Coffield et al., 2019; Mitsopoulos and Mallinis, 2017). However, these lower-accuracy works create classes based upon the ground-truth burned area size instead of clusters on the covariates as done by Shidik and Mustofa (2014).

While few works have explicitly accounted for temporal information, Liang et al. (2019) compare a recurrent neural network (RNN) and an LSTM to predict a numerical custom fire severity metric. Their results indicate that the LSTM outperforms the RNN, motivating our use of an LSTM. In addition, they discuss that their meteorological covariates are associated with fire severity, suggesting they are worth considering in the peatlands fire problem as well.

However, a two-dimensional map of predicted fire perimeters is easier for researchers, policymakers, and fire agencies to analyze and use. Recently, deep learning methods such as reinforcement learning, CNNs, and graph neural network (GNN) models have gained more attention in mapping fire spread.

Reinforcement learning models produce predicted fire perimeters by viewing the fire as an agent and modelling actions the agent is likely to take (Zheng et al., 2017; Subramanian and Crowley, 2017; Ganapathi Subramanian and Crowley, 2018). However, this formulation of the problem does not apply in the context of peatland fires since fires are no longer limited to spreading to the areas around them. Peat fires can go underground and resurface elsewhere, which is parallel to a delayed jump action for a fire agent. As such, traditional reinforcement learning algorithms would have to be altered in order to function for peatlands.

The current state-of-the-art machine learning algorithms for forest fire prediction use convolutional neural networks (CNN) (Hodges et al., 2019; Radke et al., 2019). Hodges et al. (2019) to predict future burn perimeters based on six-hourly burn maps generated by the FARSITE physics-based simulator. Radke et al. (2019) attempts to use a similar CNN architecture named FireCast based on daily observed fire perimeters from GEOMAC instead of simulation burns. FireCast is able to outperform FARSITE, which establishes some of the limitations of using simulated burns for training as compared to observed data. FireCast's performance emphasizes recall over precision with very high recall percentages but very low precision.

Few works take into account abnormally shaped spatial information for fire prediction. However, Jin et al. (2020) uses graph-convolutional layers in a custom architecture, USFP-Net. The USFP-Net uses a graph convolutional neural network, CNN layers and RNNs to model the fire prediction problem. USFP-Net outperforms many other common architectures for urban fire prediction. They represent the area as a graph with edges with weights inversely proportional to the distance between them, resulting in a fully connected graph. Training such a network would require a long time, and it is likely computationally infeasible to use this model/replicate such results over larger areas. Further, in contrast to PeatNet, the model only accounts for the usual spatial and temporal characteristics of an area without taking into account fire spread patterns.

2.2 Peatland Studies

We have surveyed literature within the general peatland domain that is potentially relevant to the problem of peatland fire prediction. Many problems arise in characterizing the peatland biome. First, it can be difficult to determine which land is peatland. DeLancey et al. (2019) predict where peatlands exist by using machine learning and boosted decision trees. Mahdianpari et al. (2018) classify wetlands as being peat-based or not by using a convolutional neural network model. Second, the peatlands have various characteristics that might be useful as covariates in fire prediction. Prior work on these elements focus on identifying the type of peatlands (Bourgeau-Chavez et al., 2017), the amount of sequestered carbon stored in the peat (Minasny et al., 2018), peatlands affected by permafrost (Hugelius et al., 2020), the acidity of the peatlands (Widyatmanti et al., 2019), identifying human draining around peatlands (Connolly and Holden, 2017), or the water table depth of the peatlands (Kalacska et al., 2018) using various remote-sensing datasets, statistical methods, or basic machine learning models.

2.3 Peatland Fire Prediction

Within the peatlands domain, there has only been a handful of studies in fire prediction, and far fewer that leverage deep-learning methods.

Honma et al. (2016); Listyorini and Rahim (2018) use a system of detectors near a specific peatland to predict fire spread. Bourgeau-Chavez et al. (2020a) perform a review of four fires in peatlands to determine how the type of a peatland affects its likelihood of burning (Bourgeau-Chavez et al., 2020b). However, these works have limited scalability since they are either restricted to areas in which they can establish a detection system or are restricted to post-event analysis.

Sitanggang et al. (2014) apply decision trees to determine active fire area, with their best model reaching an accuracy of 71.66%. However, this model does not perform any forward predictions; it predicts MODIS hot spots based upon the conditions of that very day. This work therefore shows that hot spots can be correlated with various climate and landscape information such as soil moisture, vegetation type, and precipitation.

Maulana et al. (2019) use logistic regression to predict active fire areas up to three months in advance, achieving an accuracy of 85.16%. The authors average climate and landscape conditions for four months to predict MODIS fire hot spots aggregated for one month.

Lozhkin et al. (2016) use a differential neural network model to predict carbon monoxide dispersion from peat fires near highways. This model suggests the ability of neural networks to capture peat emissions patterns.

3 Methodology

The spatiotemporal nature of the peatland fire problem lends itself well to machine learning. Ideally, a model should capture both the spatial and temporal characteristics of the peatland fires. In particular, we have tested several models including U-Net, UNet-LSTM, PT-Net, and PeatNet. In order to compare our results to previous work, we also implemented and tested a logistic regression model as a baseline. This section describes the various model architectures used for the prediction task.

3.1 Logistic Regression

We use logistic regression as our baseline classifier. The input to the classifier is a flattened feature vector for each pixel considered. Batch gradient descent is then used to learn a decision boundary for classification on a pixel level. Thus, the model is unable to properly account the spatial and the temporal characteristics of the data.

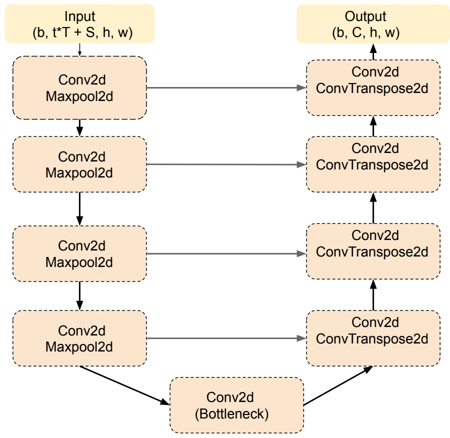

Figure 3.1: Sample U-Net architecture: Left side consists of downsampling blocks and the right side consists of upsampling blocks. The arrows represent the flow of the output. Variables: t - number of time steps, T - number of temporal features, S - number of static features, b - batch size, h - height, w - width

3.2 U-Net

U-Net has shown great promise in several image-related tasks such as object detection and segmentation (Ronneberger et al., 2015). The model consists of several downsampling blocks that create an encoding of the input variable, followed by a series of upsampling blocks that sequentially build the output. Each subsequent downsampling block uses a set of convolutional and max-pooling layers to increase the number of features and reduce the size of each channel. Each upsampling layer uses a set of convolutional layers and convolutional transpose layers to decrease the number of channels and increase the size of each channel. The downsampling layers learn increasingly coarse information about the input, which is then fed to the upsampling layers. In our implementation, different features represent different channels in the input. For features with a temporal component, such as CO_2 emissions or VIIRS, the different time-steps are regarded as different channels.

3.3 UNet-LSTM

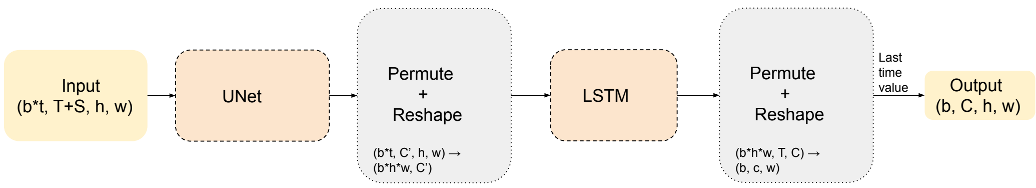

Figure 3.2: UNet-LSTM Architecture: Each batch is first fed into the U=Net component and then reshaped before being passed into the LSTM layer. b: batch size, h: height, C': number of output channels from the U-Net component, C: number of output channels–one for each class

We implement a UNet-LSTM that is able to learn both the spatial information and temporal information. The inputs are first passed through the U-Net component and then to the LSTM layer. The U-Net component considers the temporal information as a part of the batch dimension, thus only learning the spatial features. The output of this component is then reshaped such that the spatial features are a part of the batch dimension and the temporal features no longer are; this is then fed into the LSTM layer, which learns only temporal relations.

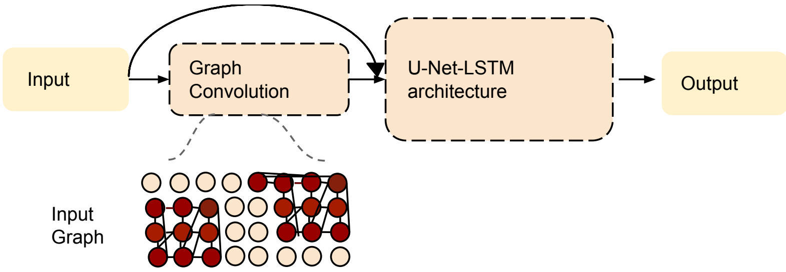

Figure 3.3: PeatNet Architecture: Brown points represent peatlands, and yellow points represent non-peatlands. Each node also has a self-loop in addition to the edges shown. Refer to the previous page for the U-Net architecture used.

3.4 PeatNet

As discussed above, peatland fires spread underground. Ideally, a model should consider that the source of the underground fire may be far away. Typical CNNs only take into account fire points that are adjacent to the current area. To account for this underground phenomenon, we construct a novel graph neural network architecture to connect 'peat points' to other non-adjacent 'peat points'. At the same time, fires are unlikely to move very far away.

Our graphical representation encodes peat points as nodes, where each peat node is connected to other peat nodes at a physical distance less than k, a hyperparameter. Further, we also add self-edges to both peat and non-peat nodes. A graph convolution is then applied to this graph, which helps the peat nodes gain additional information about other peat nodes. In particular, the graph convolution function essentially represents each node as an aggregation of its neighbors. Consequently, the encoding created for peat points is dependent on the other peat points in addition to itself, while the encoding for non-peat points is dependent on just the point itself. We then pass the final result of the graphical convolution and the original input features to the UNet-LSTM component. The output of the UNet-LSTM is finally passed through a fully connected layer to yield the final output.

3.5 PT-Net

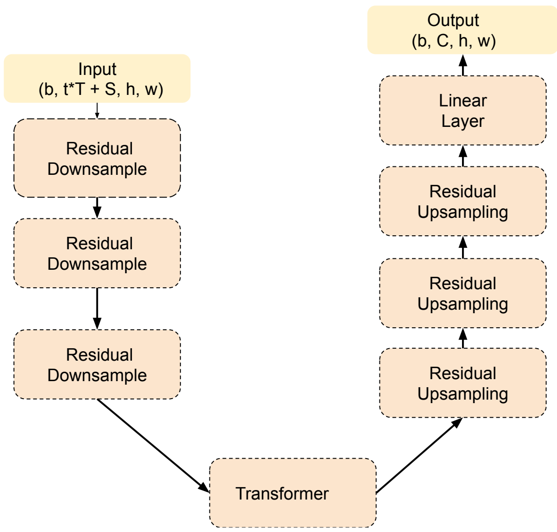

Figure 3.4: PT-Net Architecture: The input batch is encoded by the down-sampling layers, accumulated by the attention modules, and then upsampled into the output. Figure notations are the same as in Figure 3.2.

As discussed above, a good model needs to be able to correctly extract the temporal and spatial information about the data. Recent state of the art studies show that attention-based neural networks

([Vaswani et al., 2017](#)) have shown to be able to capture sequence based data better than other neural networks such as LSTMs. Accordingly, we extend upon the LSTM in modeling temporal information. Specifically, we implement a residual encoder-decoder based upon the work of He et al. (2015) with three downsampling blocks in the encoder and three upsampling blocks in the decoder. We account for temporal relations with a transformer (Vaswani et al., 2017) module, which has three multi-head self-attention layers so it can focus on multiple time-steps to predict the future time step, allowing flexible access to all the temporal features independent of length of the sequence.

4 PeatSet

Dataset	Features	Spatial Resolution	Temporal Resolution
CWFIS	`BURNCLASS`	variable	daily
CarbonTracker (Global)	`height_i` (for i in $[0,10)$)	3°x 2°	3-hourly
CarbonTracker (Flux)	`fire_flux`, `fuel_flux`	1°x 1°	3-hourly
VIIRS	`frp`, `confidence`, `bright_ti4`	375m x 375m	daily
Tarnocai Peatland Map	`TOCC`	variable	fixed
ERA5	`swvli` (for i in $[1,4]$), `stli` (for i in $[1,4]$), `lai_hv`, `lai_lv`, `tp`, `t2m`, `u10`, `v10`	0.1° x 0.1°	hourly

Table 4.1: Covariate features used for the prediction of fire. Datasets are presented along with the relevant covariate features, spatial resolution, and temporal resolution. Spatial resolution indicated with degrees is given by longitude/latitude.

Our second contribution is the curation of a collection of the first comprehensive Peat-Fire dataset consisting of previously existing remote sensing/manually labelled datasets, which we use for the tasks of predicting CWFIS burned area categories Table 4.1 presents the features used for the prediction of fires. The spatial region of our dataset covers Canada; we focus on Canada due to the relative abundance of data and area of peatlands. The time range of our dataset is from January 20, 2012 to December 31, 2018, which is the intersection of the available time ranges for the datasets. Refer to Appendix I for the exact coordinates used to specify the spatial bounds, details on each feature, and the source of each dataset.

Peatland Features: To delineate peatland from other land, we use the Tarnocai Peatland Map, which is the standard dataset used for determining where peatlands are in Canada. It is presented as a shapefile with a set of polygons. Polygons have an associated percentage of peatland cover, `PEATLAND_P`.

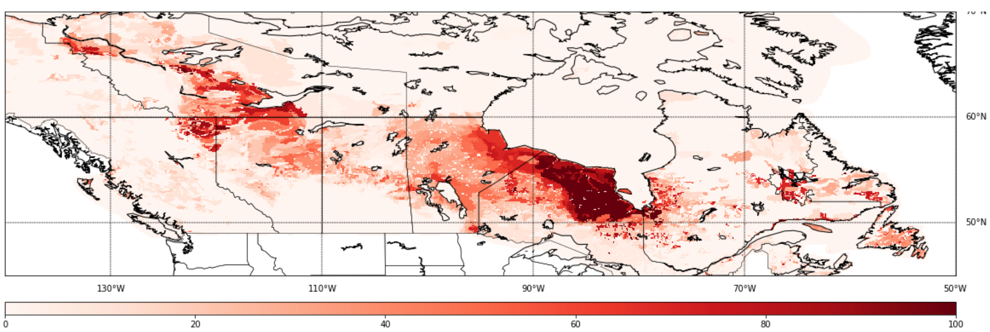

Figure 4.1: Tarnocai Peatland heatmap indicating percentage(%) of peatland cover.

Fire Features: We use the burned area product from the Canadian Wildland Fire Information System (CWFIS) (Service, b), the most comprehensive dataset for fires in Canada. This dataset is an assimilated dataset that utilizes manual reporting from governmental agencies which should help account for underground and long-lasting smouldering fires better than remote sensing sources.

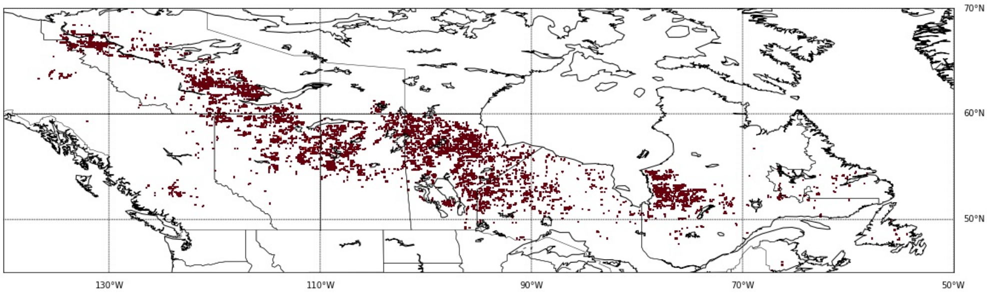

Figure 4.2: Burned area in the peatlands by masking the CWFIS burned area with Tarnocai. CWFIS data was max aggregated from 2012 to 2018.

CO_2 Emissions Features: We use the Global Monitoring Laboratory Carbon Tracker CT2019, which has a 3-by-2 longitude/latitude resolution across the globe for CO_2 emissions. CT2019 also includes the flux of CO_2 across the globe. Flux is the gradient of concentration, and determines the source of the CO_2 emissions and the cause, e.g. fire, fossil fuels. We theorize that CO_2 emissions can serve as a latent variable to represent underground fires since underground fires still output significant CO_2.

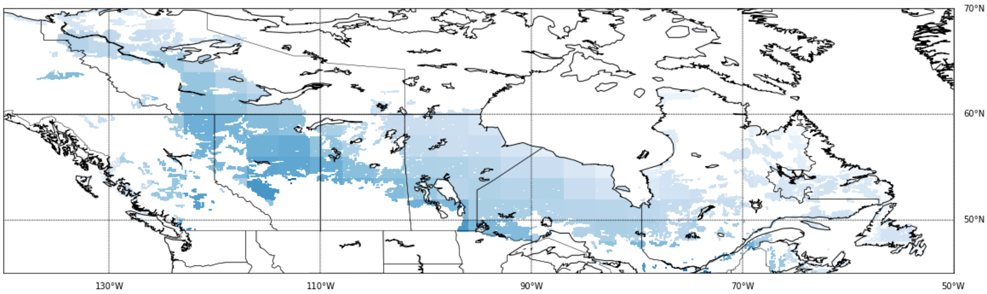

Figure 4.3: CO_2 emissions in the peatlands by masking the Global CarbonTracker data with Tarnocai. Data is in units of molar density, and was mean aggregated from 2012 to 2018.

Soil Features: We also include features pertaining to amount of carbon stored in the land, as this is a basic indicator of how much CO_2 is emitted if a fire burns over a given area. The amount of carbon stored per polygon is computed in Tarnocai using a combination of other factors, such as peat depth.

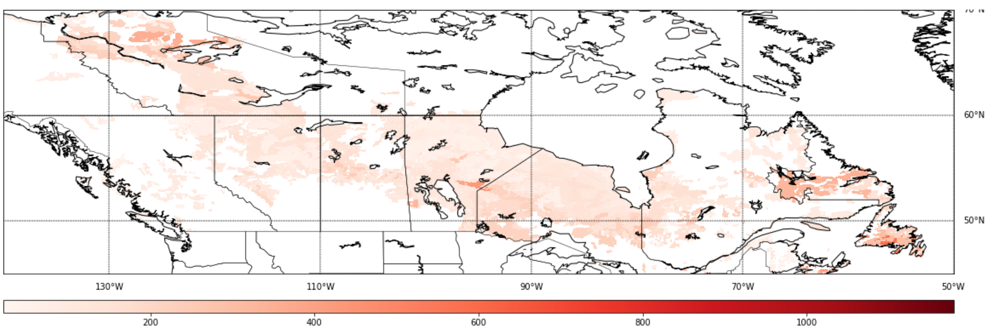

Figure 4.4: Total organic carbon content (kg/m^2) in the peatlands by the Tarnocai dataset.

Hot Spot Features: Hot spots are a good indicator of where fires are burning. We therefore incorporate hot spot data from the VIIRS dataset. Due to its finer resolution, the VIIRS satellite sensor imagery captures smaller fires than MODIS, which is the standard satellite-based data used by many fire datasets such as the Global Fire Emissions Database (GFED). In general, hot spots have a confidence of being associated with a fire, which is based on its temperature. However, peatland fires burn at a lower temperature than traditional fires. Consequently, low-confidence hot spots that persist over a long duration are potentially indicative of peatland fires.

Soil and Weather Features: We consider ERA5 which includes features pertaining to the soil and the weather. We use soil moisture and soil temperature as indicators of a fire and wind velocity to examine the spread of CO_2 from a source.

Feature Processing: The datasets are explicitly projected to WGS84 (EPSG: 4326), the standard longitude/latitude coordinate system. As each dataset has a different spatial resolution, they are all scaled to have a resolution of $0.1°$ x $0.1°$ longitude/latitude per pixel, such that each feature at a timestamp has dimensions of 483 by 910 pixels. We take daily timestamps. For features with a sub-daily resolution, we take the average over sub-daily data points. For features that do not change over time, such as `TOCC` from the Tarnocai dataset, we simply reuse the same data for each timestamp.

Additional Predicted Feature Processing: For prediction, we mask the predicted features, fire occurrence or CO_2 emissions, with the Tarnocai Peatland shapefiles to get only the values over peatlands; we do not do this for the covariates. A polygon in the Tarnocai Peatland Map is considered to be peatland if it has at least 10% peatland cover as specified by `PEATLAND_P`. A polygon in CWFIS is considered to be burned if it is estimated 100% burned as given by the `BURNCLAS` feature.

5 Results

	Recall	Precision	F1	Accuracy
Logistic Regression	0.8186	0.0016	0.0032	0.4298
UNet	0.9906	0.0212	0.0419	0.9607
UNet-LSTM	**0.9944**	0.0294	0.0571	0.9650
PeatNet	0.9668	0.0274	0.0532	0.9632
PT-Net	0.9232	**0.0532**	**0.1006**	**0.9984**

Table 5.1: Results for each of the models.

The models are evaluated on both recall and precision. Each model is provided 5 days of covariate input data and predicts CWFIS burn classes for the subsequent day. Note that for this task, recall is defined as the fraction of the fires correctly predicted over the total number of fires and precision is defined as the fraction of the correct fires predicted over the total number of fires predicted. [3]

We use binary cross entropy as the loss function for training the models. Note that the dataset is heavily skewed since there are far fewer points under fire than not under fire. Further, similar to previous studies, we emphasize recall over precision: it is more important to predict fires than to correctly predict the non-fires. We weigh the fire to non-fire class by 0.001: 1 because the non-fire class is roughly a thousand times greater than the fire class.

6 Discussion

Our results confirm that spatial information is vital for peatland fire prediction. The regression model, which does not have information from nearby blocks, performs significantly worse than the models with access to spatial information.

We also discover that access to temporal information improves peat fire prediction. The UNet-LSTM and PT-Net, which both use temporal and spatial features, outperforms the U-Net, which does not capture temporal information. PT-Net also takes advantage of both features in its architecture, outperforming PeatNet, U-Net, and UNet-LSTM.

[3] Code and dataset details available at https://github.com/Sbali11/PeatlandFirePrediction.

A likely explanation for the higher performance of UNet-LSTM as compared to PeatNet with a GNN layer is that locality generally dominates in fire modelling; therefore, the long-distance relationships the GNN models are less relevant.

7 Limitations/Future Work

7.1 Limitations

A limitation of our work was the computational challenge of training models due to the considerable scale of the data. Most models took several hours to finish one epoch of training and evaluation. Given the short time period of ProjectX, we were unable to test as many models or as many variations on hyper-parameters as we would have wished.

In addition, there are a few limitations with our dataset. There were only six years of available data across the features, which is fairly small for the types of deep learning architectures we implement. The features in our dataset are sparse, and thus the total number of interesting data points is relatively low. We also generate the dataset at a fairly low resolution, potentially erasing details that could have been relevant for the fire prediction task.

7.2 Future Work

Future work can potentially improve the quality of predictions by considering additional covariates. In this work, we focus mostly on environmental covariates; however, previously-mentioned research has established that human activity and drainage play a significant part in the damage of peatlands. Humans are also generally one of the primary causes of fire ignition. As such, integrating covariates that indicate human activity might improve predictions.

We will also want to implement more baselines to compare with our models, such as random forests, support vector machines, and vanilla CNNs.

Another area of future research could be to predict numerical fire severity metrics for each location instead of predicting whether a location is on fire or not. We already ran experiments running PT-Net for this purpose with a R^2 of 0.583. We measure fire severity by predicting CO_2 emissions, which potentially has the benefit of identifying underground fires that are not detected by VIIRS or CWFIS. In the future, we will want to implement emission prediction for our other models. We may also try to integrate different parameters of severity, such as fire radiative potential or temperature, to minimize the effect of noise within CO_2 emissions datasets.

Lastly, some of our models might prove useful in forest fire prediction as well as peat fire prediction. The success of PT-Net in particular suggests the potential of attentional models to deal with time-series fire data.

8 Conclusion

In this work, we first develop several novel architectures and adapt newer machine learning models to the problem of peat fire prediction. Our best model, PT-Net, shows great improvement over other models for fire prediction. We show experimentally that our approaches which consider the spatiotemporal aspects of the data outperform those that do not.

Second, we assemble a collection of relevant datasets to enable future studies of peatland fires. We hope that the data collection we provide will facilitate further research into peatland fire prediction.

Accurate fire spread and severity prediction can allow decision makers to invest their attention to areas with potentially severe peat fires and therefore decrease associated environmental and economic harms. The prevention of peat fires would lead to a significant decrease in global fire emissions. We hope our current work will inspire more applications of machine learning within the peatland domain to this end.

9 Acknowledgments

We would like to thank Priya Donti, David Marvin, Jan Dutton, Ann Lee, Niccolo Dalmasso, Irwin McNeely, Gordon Hamish, Xinyan Huang, Christine Weidinmyer, Laura Kiely, Amy Braverman, Sophie Wilkinson, Alexandra Barthelmes, Brian Simpson, Mikael Kuusela, and Michael Stanley for sharing their advice and knowledge during the course of the project. We would like to give special thanks to Justin Khim, our mentor, and Reid Simmons, the Director of the AI Major at CMU for bringing us together. Finally, We would like to thank the organizers of ProjectX from the University of Toronto for giving us the opportunity and impetus to develop and present this work.

References

A. R. Jacobson, K. N. Schuldt, J. B. M. T. O. P. T. A. A. J. M. L. O. G. J. C. T. A. S. A. K. A. S. A. F. A. B. B. P. B. A. B. S. C. B. A. B. D. B. G. B. J. B. A. G. C. H. C. L. C. . M. Z. (2020). Carbontracker ct2019. model published by noaa earth system research laboratory, global monitoring division, 2020.

Bourgeau-Chavez, L. L., S. Endres, R. Powell, M. J. Battaglia, B. Benscoter, M. Turetsky, E. S. Kasischke, and E. Banda (2017). Mapping boreal peatland ecosystem types from multitemporal radar and optical satellite imagery. *Canadian Journal of Forest Research 47*(4), 545–559.

Bourgeau-Chavez, L. L., S. L. Grelik, M. Billmire, L. K. Jenkins, E. S. Kasischke, and M. R. Turetsky (2020a). Assessing boreal peat fire severity and vulnerability of peatlands to early season wildland fire.

Bourgeau-Chavez, L. L., S. L. Grelik, M. Billmire, L. K. Jenkins, E. S. Kasischke, and M. R. Turetsky (2020b). Assessing boreal peat fire severity and vulnerability of peatlands to early season wildland fire.

Castelli, M., L. Vanneschi, and A. Popovič (2015). Predicting burned areas of forest fires: an artificial intelligence approach. *Fire ecology 11*(1), 106–118.

Coffield, S. R., C. A. Graff, Y. Chen, P. Smyth, E. Foufoula-Georgiou, and J. T. Randerson (2019). Machine learning to predict final fire size at the time of ignition. *International journal of wildland fire*.

Connolly, J. and N. M. Holden (2017). Detecting peatland drains with object based image analysis and geoeye-1 imagery. *Carbon balance and management 12*(1), 7.

Cortez, P. and A. d. J. R. Morais (2007). A data mining approach to predict forest fires using meteorological data.

Dadap, N. C., A. R. Cobb, A. M. Hoyt, C. F. Harvey, and A. G. Konings (2019). Satellite soil moisture observations predict burned area in southeast asian peatlands. *Environmental Research Letters 14*(9), 094014.

DeLancey, E. R., J. Kariyeva, J. T. Bried, and J. N. Hird (2019). Large-scale probabilistic identification of boreal peatlands using google earth engine, open-access satellite data, and machine learning. *Plos one 14*(6), e0218165.

Flannigan, M. D., M. A. Krawchuk, W. J. de Groot, B. M. Wotton, and L. M. Gowman (2009). Implications of changing climate for global wildland fire. *International journal of wildland fire 18*(5), 483–507.

Ganapathi Subramanian, S. and M. Crowley (2018). Using spatial reinforcement learning to build forest wildfire dynamics models from satellite images. *Frontiers in ICT 5*, 6.

Harris, N., S. Minnemeyer, F. Stolle, and O. Payne (2015, October). Indonesia's fire outbreaks producing more daily emissions than entire us economy.

He, K., X. Zhang, S. Ren, and J. Sun (2015). Deep residual learning for image recognition. corr abs/1512.03385 (2015).

Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, et al. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society 146*(730), 1999–2049.

Hodges, J. L., B. Y. Lattimer, and K. D. Luxbacher (2019). Compartment fire predictions using transpose convolutional neural networks. *Fire Safety Journal 108*, 102854.

Honma, T., K. Kaku, A. Usup, and A. Hidayat (2016). Detection and prediction systems of peat-forest fires in central kalimantan. In *Tropical Peatland Ecosystems*, pp. 397–406. Springer.

Hugelius, G., J. Loisel, S. Chadburn, R. B. Jackson, M. Jones, G. MacDonald, M. Marushchak, D. Olefeldt, M. Packalen, M. B. Siewert, et al. (2020). Large stocks of peatland carbon and nitrogen are vulnerable to permafrost thaw. *Proceedings of the National Academy of Sciences 117*(34), 20438–20446.

International Union for Conservation of Nature (2017, November). Peatlands and climate change.

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. *arXiv preprint arXiv:2003.00646*.

Jin, G., C. Zhu, X. Chen, H. Sha, X. Hu, and J. Huang (2020). Ufsp-net: a neural network with spatio-temporal information fusion for urban fire situation prediction. In *IOP Conference Series: Materials Science and Engineering*, Volume 853, pp. 012050. IOP Publishing.

Kalacska, M., J. P. Arroyo-Mora, R. J. Soffer, N. T. Roulet, T. R. Moore, E. Humphreys, G. Leblanc, O. Lucanus, and D. Inamdar (2018). Estimating peatland water table depth and net ecosystem exchange: A comparison between satellite and airborne imagery. *Remote Sensing 10*(5), 687.

Liang, H., M. Zhang, and H. Wang (2019). A neural network model for wildfire scale prediction using meteorological factors. *IEEE Access 7*, 176746–176755.

Listyorini, T. and R. Rahim (2018). A prototype fire detection implemented using the internet of things and fuzzy logic. *World Trans. Eng. Technol. Educ 16*(1), 42–46.

Lozhkin, V., D. Tarkhov, V. Timofeev, O. Lozhkina, and A. Vasilyev (2016). Differential neural network approach in information process for prediction of roadside air pollution by peat fire. In *IOP conference series: materials science and engineering*, Volume 158, pp. 012063. IOP Publishing.

Mahdianpari, M., B. Salehi, M. Rezaee, F. Mohammadimanesh, and Y. Zhang (2018). Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery. *Remote Sensing 10*(7), 1119.

Markuzon, N. and S. Kolitz (2009). Data driven approach to estimating fire danger from satellite images and weather information. In *2009 ieee applied imagery pattern recognition workshop (aipr 2009)*, pp. 1–7. IEEE.

Maulana, S. I., L. Syaufina, L. B. Prasetyo, and M. N. Aidi (2019). Spatial logistic regression models for logistic repression models for predicting peatland fire in bengkalis regency, indonesia. *Journal of Sustainability Science and Management 14*(3), 55–66.

Minasny, B., B. I. Setiawan, S. K. Saptomo, A. B. McBratney, et al. (2018). Open digital mapping as a cost-effective method for mapping peat thickness and assessing the carbon stock of tropical peatlands. *Geoderma 313*, 25–40.

Mitsopoulos, I. and G. Mallinis (2017). A data-driven approach to assess large fire size generation in greece. *Natural Hazards 88*(3), 1591–1607.

Radke, D., A. Hessler, and D. Ellsworth (2019). Firecast: Leveraging deep learning to predict wildfire spread. In *IJCAI*, pp. 4575–4581.

Ronneberger, O., P. Fischer, and T. Brox (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241. Springer.

Service, C. F. Canadian national fire database – agency fire data. Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Edmonton, Alberta.

Service, C. F. National burned area composite (nbac). Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Edmonton, Alberta.

Shidik, G. F. and K. Mustofa (2014). Predicting size of forest fire using hybrid model. In *Information and Communication Technology-EurAsia Conference*, pp. 316–327. Springer.

Short, P. Peatland distribution – in the world.

Sitanggang, I. S., R. Yaakob, N. Mustapha, and A. Ainuddin (2014). A decision tree based on spatial relationships for predicting hotspots in peatlands. *Telkomnika 12*(2), 511.

Subramanian, S. G. and M. Crowley (2017). Learning forest wildfire dynamics from satellite images using reinforcement learning. In *Conference on reinforcement learning and decision making*.

Tarnocai, C., I. M. Kettles, and B. Lacelle (2011a). Peatlands of Canada. Geological Survey of Canada.

Tarnocai, C., I. M. Kettles, and B. Lacelle (2011b). Soil Organic Carbon Content of Canadian Peatlands. Geological Survey of Canada.

Tarnocai, C., I. M. Kettles, and B. Lacelle (2011c). Soil Organic Carbon Mass of Canadian Peatlands. Geological Survey of Canada.

Thompson, D. (2020, July). Peatland fires and carbon emissions.

Turetsky, M., K. Wieder, L. Halsey, and D. Vitt (2002). Current disturbance and the diminishing peatland carbon sink. *Geophysical Research Letters 29*(11), 21–1.

Turetsky, M. R., B. Benscoter, S. Page, G. Rein, G. R. Van Der Werf, and A. Watts (2015). Global vulnerability of peatlands to fire and carbon loss. *Nature Geoscience 8*(1), 11–14.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin (2017). Attention is all you need. In *Advances in neural information processing systems*, pp. 5998–6008.

Widyatmanti, W., D. Umarhadi, M. Ningam, Z. Sarah, K. Nugroho, Y. Sulaeman, et al. (2019). Mapping acid sulfate soil hydrogeomorphical unit on the peatland landscape using a hybrid remote sensing approach. In *Tropical Wetlands-Innovation in Mapping and Management: Proceedings of the International Workshop on Tropical Wetlands: Innovation in Mapping and Management, October 19-20, 2018, Banjarmasin, Indonesia*, pp. 30. CRC Press.

Wieder, R. K. and D. H. Vitt (2006). *Boreal peatland ecosystems*, Volume 188. Springer Science & Business Media.

Wiedinmyer, C., S. Akagi, R. J. Yokelson, L. Emmons, J. Al-Saadi, J. Orlando, and A. Soja (2011). The fire inventory from ncar (finn): A high resolution global model to estimate the emissions from open burning. *Geoscientific Model Development 4*(3), 625.

Zheng, Z., W. Huang, S. Li, and Y. Zeng (2017). Forest fire spread simulating model using cellular automaton with extreme learning machine. *Ecological Modelling 348*, 33–43.

Appendix I: Dataset Annotations

We list features we considered interesting from the following datasets, whether they were eventually used or not. For the purpose of our research, the features used were bound by the following South/North/West/East coordinates: -141.0000°/-50.0000°/41.7500°/90.0000°.

1. Tarnocai Peatland Map (Tarnocai et al., 2011a)(Tarnocai et al., 2011b)(Tarnocai et al., 2011c)
   A set of shapefiles capturing where the peatlands are, what type (bog, fen, swamp marsh) they are, and how much carbon they store. The data was gathered through survey and published by National Resources Canada.
   a) Spatial Range: Canada
   b) Variables:
      i. Peatland (%) (`PEATLAND_P`): Percentage of shapefile polygon covered in peatland.
      ii. `BOG_PCT`: Percentage of shapefile polygon covered in bog.
      iii. `FEN_PCT`: Percentage of shapefile polygon covered in fen.
      iv. `SWAMP_PCT`: Percentage of shapefile polygon covered in swap.
      v. `MARSH_PCT`: Percentage of shapefile polygon covered in marsh.
      vi. `TOCC`: The average amount of carbon stored per surface area (kg/m^2) for the shapefile polygon (uses peat depth).
2. Canadian Wildland Fire Information System (CWFIS) (Service, b)(Service, a)
   Shapefiles indicating where fires occurred and their burned areas. Data on fires was collected through survey by Canadian fire management agencies. Data on burned area was calculated through a combination of survey data, aerial photography, and satellite data, such as from through MODIS, VIIRS, Landsat, and Sentinel-2.
   a) Spatial Range: Canada
   b) Temporal Resolution: Daily
   c) Temporal Range: < January 2000 - January 2019
   d) Variables:
      i. `BURNCLAS`: Proportion of land burned for shapefile polygon. 1: estimated 25% burned, 2: estimated 50% burned, 3: estimated 75% burned, 4: estimated 100% burned.
      ii. Fire: Shapefile polygons outlining fires.
3. VIIRS
   Location of hot spots, which are thermal anomalies that often indicate fire. The data is gathered by the VIIRS sensor, onboard the Suomi NPP and NOAA-20 polar-orbiting satellites.
   a) Spatial Range: Global
   b) Spatial Resolution: 375m x 375m
   c) Temporal Range: January 2012 - present
   d) Temporal Resolution: Daily
   e) Variables:
      i. `frp`: Fire radiative power, megawatts.
      ii. `confidence`: Confidence of individual hot spots. 0: low, 1: nominal, 2: high.
      iii. `bright_ti4`: Fire pixel channel I4 brightness temperature (Kelvin).
4. Global Monitoring Laboratory Carbon Tracker CT2019 Globe (A. R. Jacobson, 2020)
   a) Spatial range: Global (NOTE: During winter time, high latitude regions are not reliable.)
   b) Spatial resolution: 3° longitude x 2° latitude
   c) Temporal Range : January 2000 - present
   d) Temporal Resolution: 3-hourly
   e) Variables:
      i. `height_i`, for $i \in [0, 10)$: CO_2 molar density at 10 different height levels above the ground. Refer to Table 1 in Section 6.1 on their documentation for the actual heights in meters.

5. Global Monitoring Laboratory Carbon Tracker CT2019 Flux (A. R. Jacobson, 2020)
   a) Spatial range: Global (NOTE: During winter time, high latitude regions are not reliable.)
   b) Spatial resolution: 1° x 1°
   c) Temporal Range: January 2000 - present
   d) Temporal Resolution: 3-hourly
   e) Variables:
      i. `fire_flux`: Flux of CO_2 attributed to fire.
      ii. `fuel_flux`: Flux of CO_2 attributed to burning fossil fuel.
6. ERA5 (Hersbach et al., 2020)
   The standard dataset on weather variables, such as pertaining to soil, precipitation, temperature, and wind. The data is an assimilation between observations and modelling of climate.
   a) Spatial Resolution: 9km x 9km regridded for 0.1° x 0.1°
   b) Temporal Resolution: hourly
   c) Temporal Range: January 1981 - 3 months before present
   d) Variables:
      i. `swvl1`: Soil water level 1. Meters-cubed water in meters-cubed soil at a depth of 0 - 7 cm from surface.
      ii. `swvl2`: Soil water level 2. Meters-cubed water in meters-cubed soil at a depth of 7 - 28 cm from surface.
      iii. `swvl3`: Soil water level 3. Meters-cubed water in meters-cubed soil at a depth of 28 - 100 cm from surface.
      iv. `swvl4`: Soil water level 4. Meters-cubed water in meters-cubed soil at a depth of 100 - 289 cm from surface.
      v. `stl1`: Soil temperature level 1. Temperature of soil in Kelvins at a depth of 0 - 7 cm from surface.
      vi. `stl2`: Soil temperature level 2. Temperature of soil in Kelvins at a depth of 7 - 28 cm from surface.
      vii. `stl3`: Soil temperature level 3. Temperature of soil in Kelvins at a depth of 28 - 100 cm from surface.
      viii. `stl4`: Soil temperature level 4. Temperature of soil in Kelvins at a depth of 100 - 289 cm from surface.
      ix. `lai_hv`: Leaf area index, low vegetation. Surface area of low-lying leaves in meters-squared over area over area land in meters-squared. Characterizes the density of low vegetation such as crops, marshes, grasses, bogs.
      x. `lai_lv`: Leaf area index, high vegetation. Surface area of high-reaching leaves in meters-squared over area land in meters-squared. Characterizes the density of high vegetation such as forests and trees.
      xi. `tp`: Total precipitation. Total amount of water accumulated over an hour. Given as the depth in meters the water would have been if spread evenly over the spatial unit. CAUTION: This variable is an aggregation and not an average, so its value describes a very specific space and time.
      xii. `t2m`: 2-meter temperature. Temperature (K) two meters above the surface.
      xiii. `u10`: East-ward component of wind. Positive-x component of speed (m/s) of wind ten-meters above surface.
      xiv. `v10`: North-ward component of wind. Positive-y component of speed (m/s) of wind ten-meters above surface.

Appendix II: Acknowledgments

Many people that we have reached out to have helped us on this project. We list them here in the order we contacted them along with their main contributions.

Justin Khim, our mentor and postdoctoral researcher in the Machine Learning Department of CMU. He was instrumental in providing us insightful feedback about identifying a meaningful problem, working as a team, and revising our paper.

Reid Simmons, robotics professor and director of the artificial intelligence major at CMU. He gathered the team, helped us meet and work together over the course of the project, and was our contact point for accessing resources from CMU, such as storage hardware for our data.

Priya Donti, Ph.D. student of computer science and public policy at CMU. She helped us define our problem space as we were exploring ideas.

David Marvin, CEO of Salo Sciences. He gave us advice about research question design and feasibility. Through our conversation with him, we were able to narrow down our focus on peatlands to fire behavior prediction as well as identify challenges in spread modeling. Spread modeling requires much higher temporal granularity in data than what is readily available. Therefore, attaining the data and developing a successful pipeline for this task would be difficult in the time constraints. Additionally, hazard modeling with ML is an area of high interest, as ML can allow for solving the problem without the need for extensive computational resources, as is the case for the physical models.

Jan Dutton, CEO of Prescient Weather, who works on the cutting-edge of climate prediction. Dutton is extremely knowledgeable about existing climate problems and had some excellent ideas on the application of ML on current climate problems. In particular, he mentioned project ideas involving extreme weather prediction and forecasting that would be vital as the frequency and power of extreme weather events continue to intensify. However, he was intrigued by our peatland fires problem after we showed him the impact it could make on carbon emissions. He mentioned many potential climate datasets we could use, and helpfully narrowed down our options to ones that were well-maintained and fairly straightforward to access given our limited time and potentially limited compute. Our conversation with him underscored the value and novelty of our peatlands research question, but also broadened our perspective on the flexibility and power of climate data in predicting a wide variety of global phenomena.

Ann Lee, head of the Statistical Methods for the Physical Sciences group at CMU, and her Ph.D. students, **Niccolo Dalmasso** and **Irwin McNeely**. After we brought up our interest in the peatland fires problem, she helped us specify the output of our problem to how fires affect CO_2 emissions. Dalmasso advised us to focus on a highly specific output, warning us that processing of the data itself for some purpose would be very time-consuming. McNeely suggested we incorporate human land-use data. Lee connected us with other CMU researchers, such as Gordon Hamish.

Gordon Hamish, research professor at CMU with a focus on air-borne pollution. He advised us to look more carefully into the GFED data, because it might be ill-posed to helping us identify fires most relevant to peatlands. The GFED data uses the MODIS satellite burned area product for fire–however, MODIS has difficulty identifying small fires and performs poorly when the ground is covered by cloud, fog, or smoke. Instead, he encouraged us to consider the more niche problem of mapping small, underground fires using field data as the ground truth. He suggested we contact Robert Yokelson, an expert on peatlands and emissions.

Xinyan Huang, professor at Hong Kong Polytechnic University who wrote his thesis on the combustion of peatlands. He helped us understand the difference between underground and aboveground peatland fires. Surface fires release more CO_2 because their intake of oxygen is faster. Underground fires burn less; however, they can last through winters to burn for many years. As there are less emissions attributed to fire during winter, we may get decent estimates of the baseline CO_2 in the region. Peat fires are distinguished by their chemical signature from forest fires–the signature of the former contains ammonia, while for the latter, it contains nitrous oxide. The ratio of carbon monoxide relative to carbon dioxide can indicate how imperfect the combustion is–the higher the ratio, the less perfect, and the more likely the fire is to be underground.

Christine Weidinmyer, professor at the University of Colorado Boulder who created the Fire INventory from NCAR (FINN) dataset, which records the modelled emissions and movement of different chemicals around the world. Most notably, she recommended we use the VIIRS dataset for hot spot detection and the MOPITT dataset for carbon monoxide detection. She also introduced us to her doctoral student, Laura Kiely.

Laura Kiely, PhD student of the University of Leeds who specializes in Indonesian peatlands. She recommended we approach analyzing emissions from a “bottom-up" perspective, that is, predict how much carbon dioxide emissions come from fire by examining the carbon content of the land. This is in contrast to observing the concentrations of carbon dioxide through satellites. The carbon dioxide disperses quickly from the source. Aas satellites have long update times (more than a few weeks), it is hard to determine if the carbon dioxide at a site came from the site or was dispersed from another site. She also informed us that field data for emissions from Indonesian peatland fires was difficult to come by for logistical and political reasons. Without field data, it is difficult to validate a model that predicts how much emissions are produced in an area.

Amy Braverman, JPL scientist with expertise in remote sensing. She told us the importance of determining whether a data source was observational, model-generated, or an assimilated. Data sources that are more observational are closer to ground truth; however, they are often noisier. She recommended we examine the NASA OCO-2 dataset for carbon dioxide. However, as the OCO-2 satellite has a polar vs. equatorial orbit, it is better at capturing data for sites further from the equator. In addition, clouds at these off-equator latitudes distort the satellite imagery less.

Sophie Wilkinson, postdoctoral fellow at McMaster University who is expert on Canadian peatlands. She helped us decide to study Canadian peatlands over Indonesian peatlands by recommending us a wealth of data associated with the former. In particular, she pointed us to the Canadian Wildland Fire Information System (CWFIS), which contains historical information spanning several decades on fires that have occurred throughout Canada. She also recommended using hot spots to identify peatland fires that emitted carbon dioxide. She also referred us to the Tarnocai peatland dataset, which is a shapefile of where peatlands are in Canada.

Alexandra Barthelmes, coordinator of the Global Peatlands Database who graciously allowed us access to the dataset.

Brian Simpson, analyst and modeller of the Canadian Forest service who answered our many questions about the burned area and fire polygon CWFIS, datasets and allowed us access to their fire intensity dataset.

Mikael Kuusela, professor of statistics and data science at CMU and his PhD student, **Michael Stanley**. Both of them are heavily involved in the OCO-2 carbon dioxide dataset. After we examined OCO-2, we realized its coverage and temporal resolution were not suitable for our needs. They recommended we look at more processed, model-based products, the most notable being the Global Monitoring Laboratory Carbon Tracker. They also suggested we consider the flux of carbon dioxide as well as the concentration to determine where carbon dioxide comes from. The Carbon Tracker is partially based on satellite data. Satellite data uses the sun to determine where carbon dioxide is. Therefore, for high northern latitudes, data captured during winter is not reliable because of the lack of sun.

ProjectX, a competition on applying machine learning to the critical issue of climate change and founded by members of the University of Toronto community. We are grateful for the opportunity given by this program to connect with other researchers, get feedback on our project, and present our work.