

## Supplementary Material

# Insights into boreal forest degradation from a canopy stability index

Brendan Mackey<sup>1\*</sup>, Sonia Hugh<sup>1</sup>, Pat Norman<sup>1</sup>, Brendan M. Rogers<sup>3</sup> Dominick Dellasala<sup>2</sup>

<sup>1</sup>Griffith University

<sup>2</sup>Woodwell Climate Research Centre

<sup>3</sup>Wild Heritage, A Project of Earth Island Institute

\*Correspondence: e-mail@e-mail.com; Tel.: (optional; include country code; if there are multiple corresponding authors, add author initials)

### Table of Contents

#### SM1 Data and methods

Figure SM1 Severe insect infestation for Ontario and Quebec

Table SM1 SRTM regrouping into simplified categories

Table SM2: Combining insect disturbance - ranking for QC

Table SM3: Combining insect disturbance - ranking for ON

Table SM4: Simplified surficial geology

#### SM2 Results

Table SM5: GBM model optimal hyperparameters

##### SM2.1 All forest – all drivers

##### SM2.2 All forest – environmental drivers

Figure SM2 Relative influence of each variable on instability for all forests – all drivers. Relative influence is a measure indicating the relative importance of each variable in training the GBM model.

Figure SM3 Partial dependence and ICE plots for all forests – all drivers

Figure SM4 Ablation study for all forests – all drivers

Figure SM5 Relative influence of each variable on instability for all forests – environmental drivers

Figure SM6 Partial dependence and ICE plots for all forests – environmental drivers

Figure SM7 Ablation study for all forests – environmental drivers

##### SM2.3 Boreal forests - environmental drivers only

Figure SM8 Relative influence of each variable on instability for boreal forests – environmental drivers

Figure SM9. Partial dependence and ICE plots for boreal forest – environmental drivers

Figure SM10. Ablation study for boreal forest – environmental drivers

## SM1 Data and methods

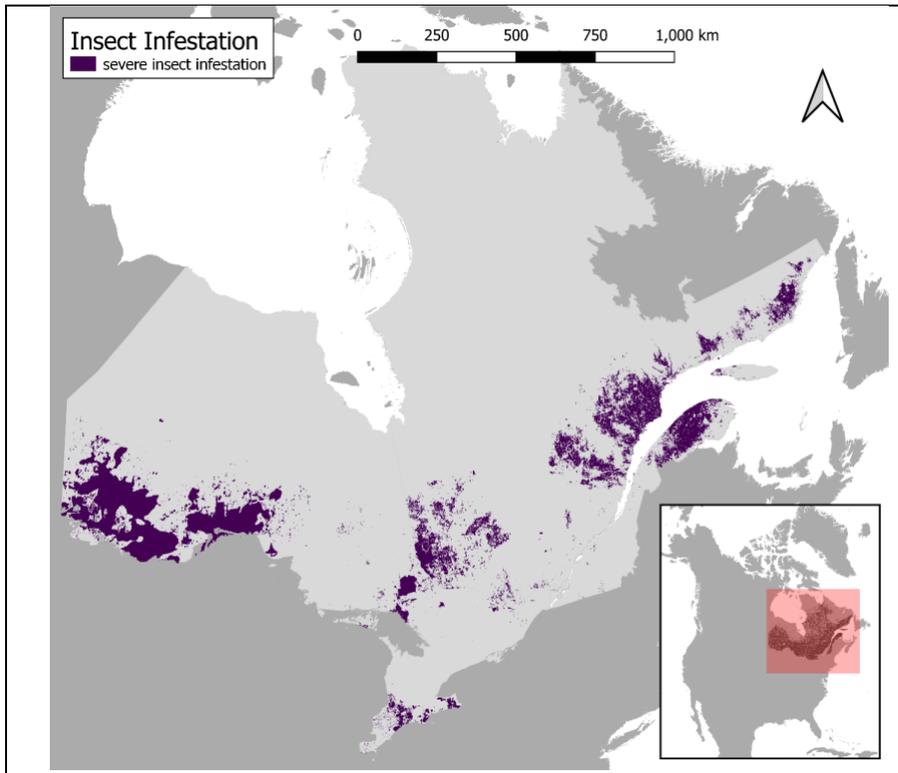


Figure SM1 Severe insect infestation for Ontario and Quebec

Table SM1 SRTM regrouping into simplified categories

Value	Simplified group	Original description
11	peak	Peak/ridge (warm)
12	peak	Peak/ridge
13	peak	Peak/ridge (cool)
14	mountain/divide	Mountain/divide
15	cliff	Cliff
21	upper slope	Upper slope (warm)
22	upper slope	Upper slope
23	upper slope	Upper slope (cool)
24	upper slope	Upper slope (flat)
31	lower slope	Lower slope (warm)
32	lower slope	Lower slope
33	lower slope	Lower slope (cool)
34	lower slope	Lower slope (flat)
41	valley	Valley
42	valley	Valley (narrow)

Table SM2: Combining insect disturbance - ranking for QC

<b>NIVEAU</b>	<b>Description</b>	<b>rank</b>
1	Light defoliation	1
2	Moderate defoliation	3
3	Severe defoliation	5
8	Forest intervention, total cutting	99
54	Mortality from 1 to 25%	6
55	Mortality from 26 to 50%	6
56	51-75% mortality	6
57	Mortality: 76 to 100%	6
99	Damage observed using satellite imagery	4

Table SM3: Combining insect disturbance - ranking for ON

<b>RANKING</b>	<b>rank</b>
Light	1
Light-Moderate	2
Moderate	3
Moderate-Severe	4
Severe	5
Mortality	6
Unknown	0
General Damage	4

Table SM4: Simplified surficial geology

<b>label</b>	<b>geology</b>	<b>relabel</b>
A	Alluvial sediments, undifferentiated	A
GFc	Glaciomarine: Ice-contact sediments	G
GFp	Glaciomarine: Outwash plain sediments	G
GLn	Glaciomarine: Littoral and nearshore sediments	G
GLo	Glaciomarine: Offshore sediments	G
GMn	Glaciomarine: Littoral and nearshore sediments	G
GMv	Glaciomarine: Veneer	G
Ln	Lacustrine: Littoral and nearshore sediments	L
Lo	Lacustrine: Offshore sediments	L

Mn	Marine: Littoral and nearshore sediments	M
O	Organic deposits, undifferentiated	O
R	Bedrock, undifferentiated	R
Tb	Glacial Till: Till blanket	T
Tm	Glacial Till: Moraine complex	T
Tv	Glacial Till: Till veneer	T

## SM2 Results

Table SM5: GBM model optimal hyperparameters

Run		Trees	Interaction depth	Minobsinnode	Shrinkage	Bag fraction	Train fraction	Cv folds	RMSE
All forest	All drivers	600	5	15	0.05	1	0.5	5	2.111
All forest	Environmental drivers	200	3	15	0.1	1	0.5	5	2.293
Boreal forest	All drivers	400	5	15	0.05	1	0.5	5	2.063
Boreal forest	Environmental drivers	200	3	5	0.05	0.5	0.5	5	2.261

### SM2.1 All forest – all drivers

The relative influence of each independent variable on the stability of all forest was calculated and plotted. The results are shown in the plot and table in Figure SM2. The four most influential variables were: insects, fire, harvest and geology. The partial dependence plots are given in Figure SM3 and the ablation study values in Figure SM4. Any of the RMSE values that are higher than the original model (the all row) performs worse when the variable is removed from the model.

### SM2.2 All forest – environmental drivers

The relative influence of each independent variable on the stability of all forest was calculated and plotted excluding fire, insect and harvesting. All three: geology, TWI and landform were influential (as seen in Figure SM5). The partial dependence plots are given in Figure SM6 and the ablation study values in Figure SM7. Any of the RMSE values that are higher than the original model (the all row) performs worse when the variable is removed from the model.

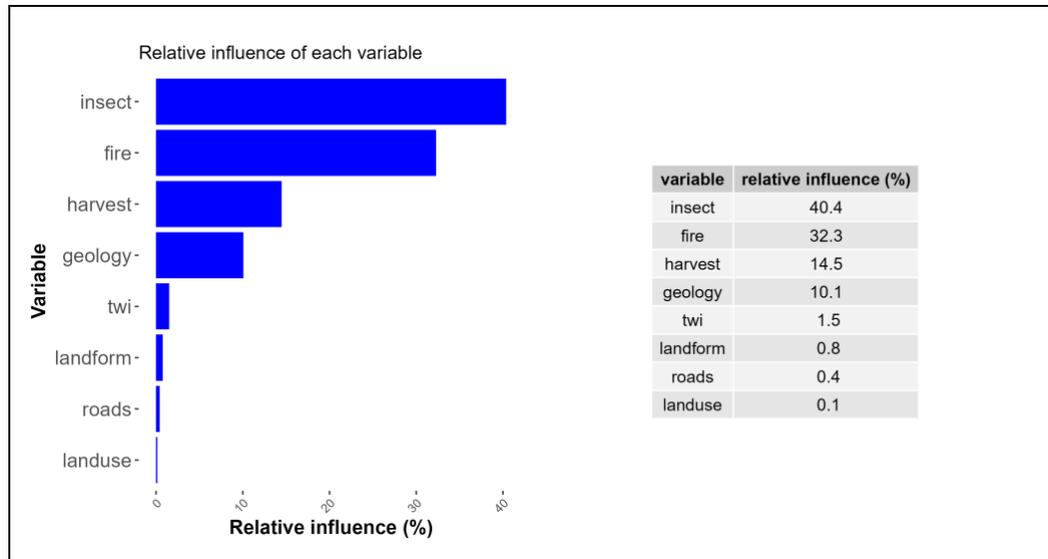


Figure SM2 Relative influence of each variable on instability for all forests – all drivers. Relative influence is a measure indicating the relative importance of each variable in training the GBM model.

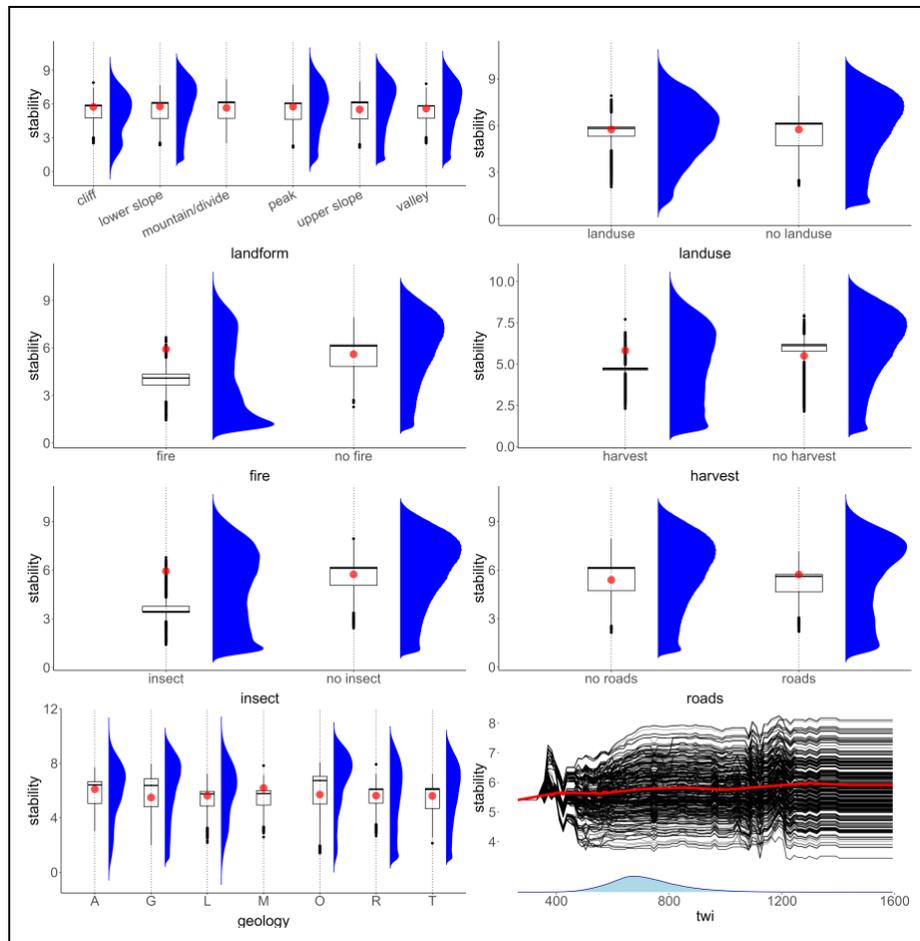


Figure SM3 Partial dependence and ICE plots for all forests – all drivers. The boxplot and black lines for TWI show the ICE values, the distribution of the predicted response variable for each observation

as we vary each predictor variable in the model. For the TWI ICE values, the values are centered to the first point of the PDP value. The red dot and red line for TWI representing the PDP value and the half violin plot (blue) shows the density of the training data.

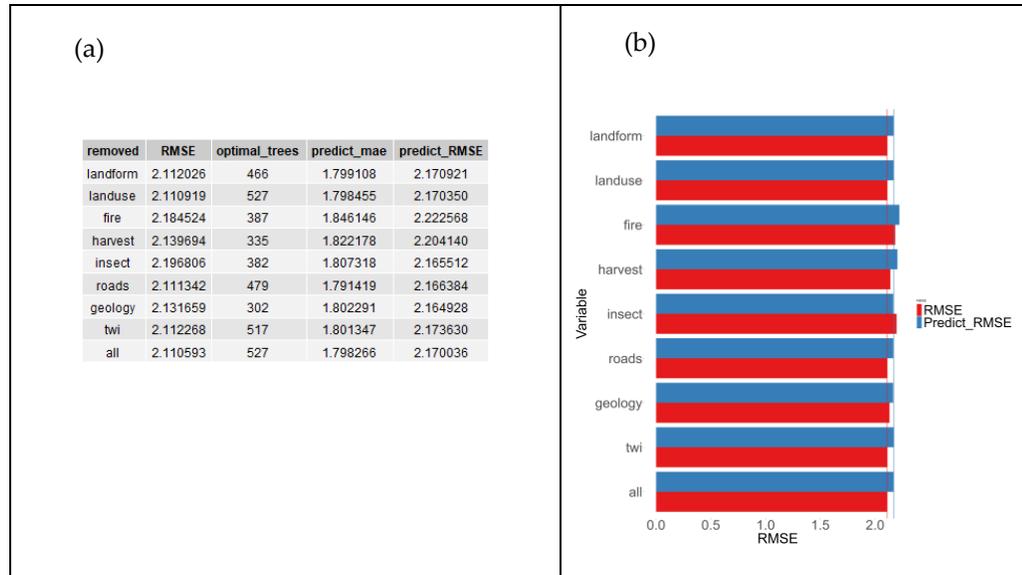


Figure SM4 Ablation study for all forests – all drivers. (a) RMSE values for the trained data and the predict RMSE values with the test data. (b) bar graph representation of the ablation RMSE values the vertical line corresponds with the RMSE values of the original model without variable removal. Any values past the vertical line, with corresponding colour, performed worse when removed.

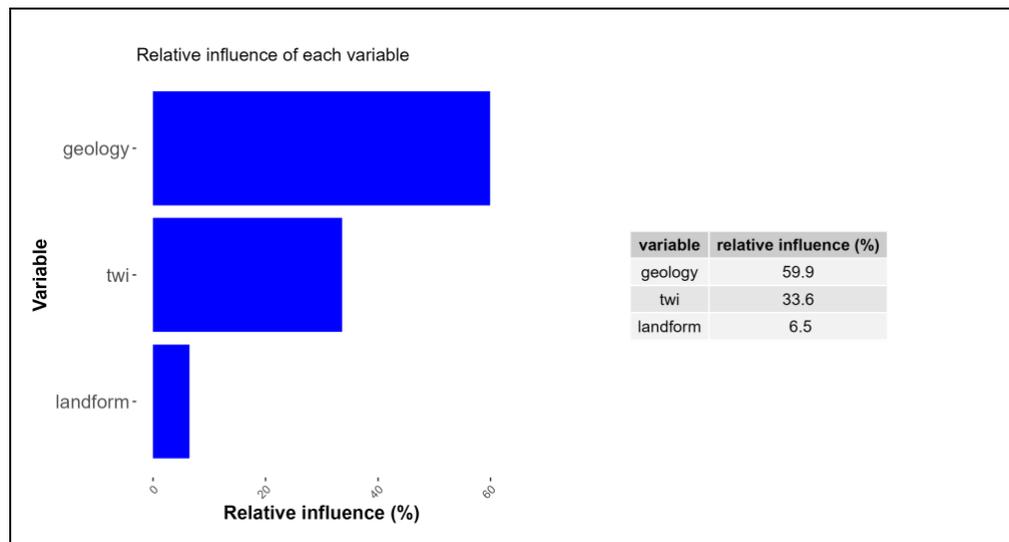


Figure SM5 Relative influence of each variable on instability for all forests – environmental drivers. Relative influence is a measure indicating the relative importance of each variable in training the GBM model.

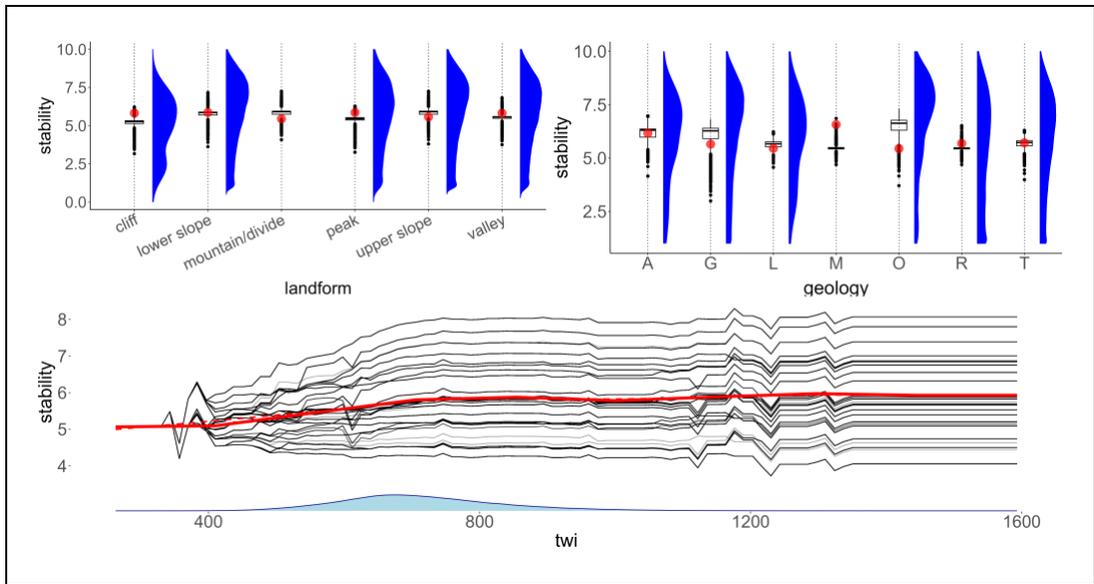


Figure SM6 Partial dependence and ICE plots for all forests – environmental drivers. The boxplot and black lines for TWI show the ICE values, the distribution of the predicted response variable for each observation as we vary each predictor variable in the model. For the TWI ICE values, the values are centered to the first point of the PDP value. The red dot and red line for TWI representing the PDP value and the half violin plot (blue) shows the density of the training data.

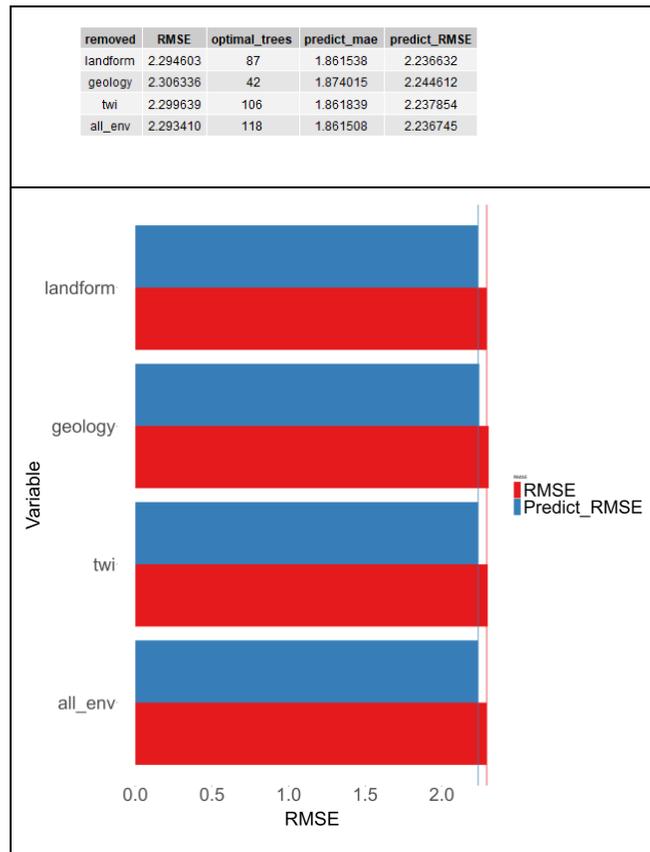


Figure SM7 Ablation study for all forests – environmental drivers (a) RMSE values for the trained data and the predict RMSE values with the test data. (b) bar graph representation of the ablation RMSE

values the vertical line corresponds with the RMSE values of the original model without variable removal. Any values after the vertical line, with corresponding colour, performed worse when removed.

### SM2.3 Boreal forests - environmental drivers only

The relative influence of each independent variable on the stability of the forest was calculated and plotted. The results are shown in the table below and the plot in Figure SM8. All three parameters have an influence: TWI, geology and landform. Partial dependence plots can be found in Figure SM9. The ablation study values can be found in Figure SM10. Any of the RMSE values that are lower than the original model (the all row) performs better when the variable is removed from the model.

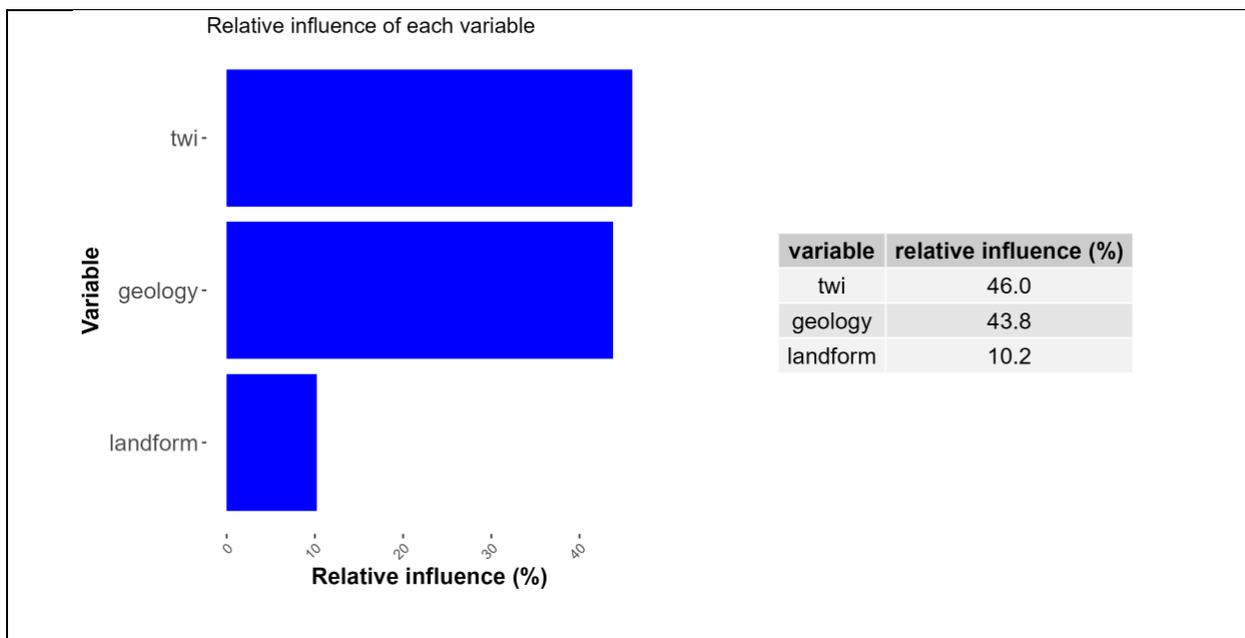


Figure SM8 Relative influence of each variable on instability for boreal forests – environmental drivers. Relative influence is a measure indicating the relative importance of each variable in training the GBM model.

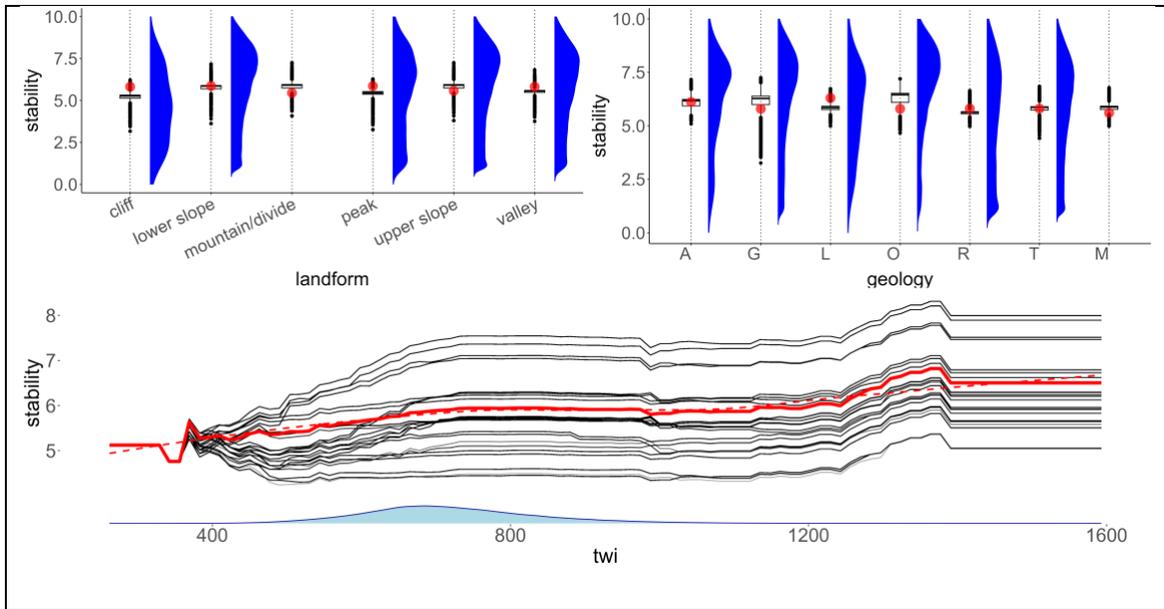


Figure SM9. Partial dependence and ICE plots for boreal forest – environmental drivers. The boxplot and black lines for TWI show the ICE values, the distribution of the predicted response variable for each observation as we vary each predictor variable in the model. For the TWI ICE values, the values are centered to the first point of the PDP value. The red dot and red line for TWI representing the PDP value and the half violin plot (blue) shows the density of the training data.

(a)

removed	RMSE	optimal_trees	predict_mae	predict_RMSE
landform	2.262921	167	1.925617	2.309386
geology	2.269401	97	1.940385	2.319263
twi	2.267688	198	1.928926	2.312904
all_boreal_env	2.261446	186	1.925094	2.309487

(b)

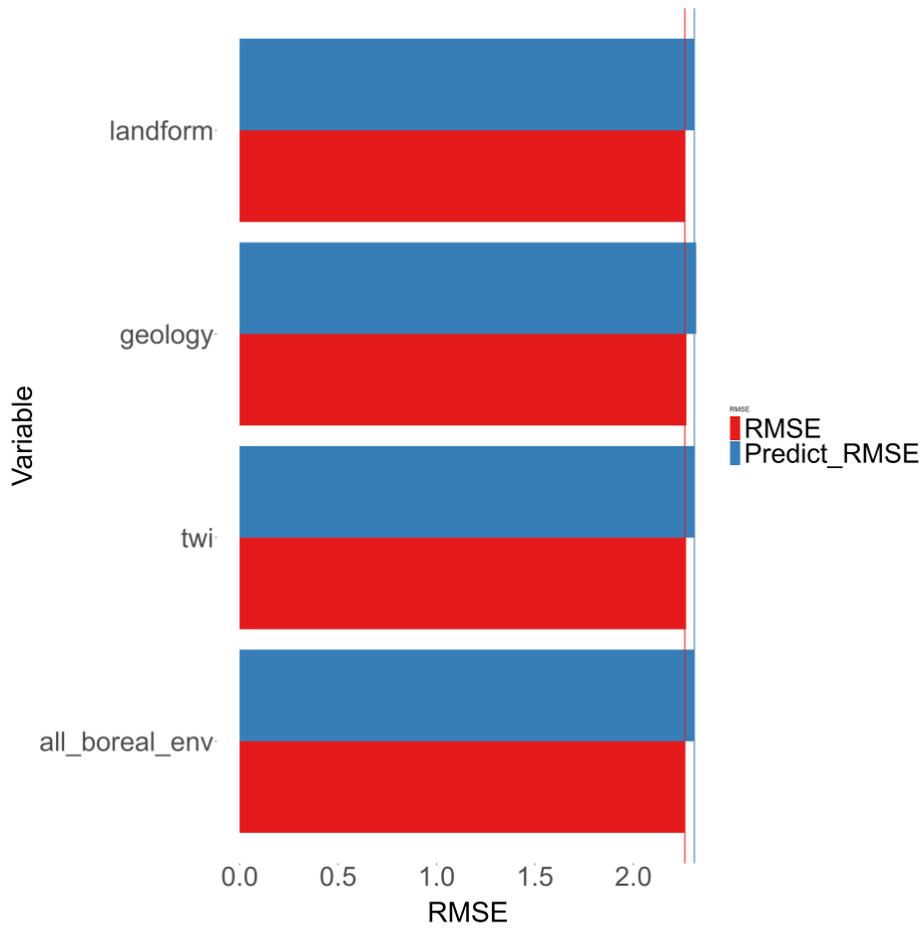


Figure SM10. Ablation study for boreal forest – environmental drivers. (a) RMSE values for the trained data and the predict RMSE values with the test data. (b) bar graph representation of the ablation RMSE values the vertical line corresponds with the RMSE values of the original model without variable removal. Any values after the vertical line, with corresponding colour, performed worse when removed.

## References

- Canadian Forest Service. 2021. Canadian National Fire Database (CNFDB). Natural Resources Canada. <https://cwfis.cfs.nrcan.gc.ca/ha/nfdb> (accessed on June 23, 2023)
- Esri, 2023. ArcGIS Pro, 3.2.1 [Computer software]. Redlands, CA: Environmental Systems Research Institute.
- GRASS Development Team, 2023. Geographic Resources Analysis Support System (GRASS) Software, Version 8.2.1. Open Source Geospatial Foundation. <https://grass.osgeo.org>
- Hermosilla, T., M.A. Wulder, J.C. White, N.C. Coops, G.W. Hobart, L.B. Campbell, 2016. Mass data processing of time series Landsat imagery: pixels to data products for forest monitoring. International Journal of Digital Earth 9(11), 1035-1054 <https://doi.org/10.1080/17538947.2016.1187673>
- Land Information Ontario. MNRF Road Segments; Land Information Ontario: Ontario, ON, Canada, 2020. Available online: <https://geohub.lio.gov.on.ca/datasets/lio::mnrf-road-segments/explore> (accessed on 17 April 2023).
- Land Information Ontario. Ontario Land Cover Compilation v.2.0; Land Information Ontario: Ontario, ON, Canada, 2022. Available online: <https://geohub.lio.gov.on.ca/documents/7aa998fdf100434da27a41f1c637382c> (accessed on 13 April 2023).
- Land Information Ontario. Forest Resources Inventory Packaged Products—Version 2; Land Information Ontario: Ontario, ON, Canada, 2022. Available online: <https://geohub.lio.gov.on.ca/maps/lio::forest-resources-inventory-packaged-products-version2/about> (accessed on 4 April 2023).
- Ministère de l'Environnement. Utilisation du Territoire—Données Québec; Ministère de l'Environnement: Québec, QU, Canada, 2023. Available online: <https://www.donneesquebec.ca/recherche/dataset/utilisation-du-territoire> (accessed on 13 April 2023).
- Ministère des Ressources naturelles et des Forêts. Feux de forêt, [Dataset], in Data Quebec, 2017, updated 2022. [<https://www.donneesquebec.ca/recherche/dataset/feux-de-foret>], (downloaded on 21 June 2023).
- Ministère des Ressources Naturelles et des Forêts. Carte Écoforestière à Jour—Données Québec; Ministère des Ressources Naturelles et des Forêts: Québec, QU, Canada, 2022. Available online: <https://www.donneesquebec.ca/recherche/dataset/carte-ecoforestiereavec-perturbations#> (accessed on 15 June 2023).

Ontario Ministry of Natural Resources and Forestry. 2020. Fire disturbance area. <https://geohub.lio.gov.on.ca/datasets/lio::fire-disturbance-area/about>

Statistics Canada. National Road Network—NRN—GeoBase Series—Open Government Portal; Statistics Canada: Ottawa, ON, Canada, 2022. Available online: <https://open.canada.ca/data/en/dataset/3d282116-e556-400c-9306-ca1a3cada77f> (accessed on 17 April 2023).

Theobald, D. M., Harrison-Atlas, D., Monahan, W. B., & Albano, C. M. (2015). Ecologically-relevant maps of landforms and physiographic diversity for climate adaptation planning. *PloS one*, 10(12), e0143619

Land Information Ontario. Forest Resources Inventory Packaged Products—Version 2; Land Information Ontario: Ontario, ON, Canada, 2022. Available online: <https://geohub.lio.gov.on.ca/maps/lio::forest-resources-inventory-packaged-products-version2/about> (accessed on 4 April 2023).

Ministère des Ressources Naturelles et des Forêts. Unité D'Aménagement (UA)—Données Québec; Ministère des Ressources Naturelles et des Forêts: Québec, QU, Canada, 2023. Available online: <https://www.donneesquebec.ca/recherche/dataset/united-amenagement> (accessed on 18 April 2023).

Ministère des Ressources Naturelles et des Forêts. Récolte et Autres Interventions Sylvicoles—Données Québec; Ministère des Ressources Naturelles et des Forêts: Québec, QU, Canada, 2023. Available online: <https://www.donneesquebec.ca/recherche/dataset/recolte-et-reboisement> (accessed on 4 April 2023).

Mackey, Brendan, Carly Campbell, Patrick Norman, Sonia Hugh, Dominick A. DellaSala, Jay R. Malcolm, Mélanie Desrochers, and Pierre Drapeau. 2024. "Assessing the Cumulative Impacts of Forest Management on Forest Age Structure Development and Woodland Caribou Habitat in Boreal Landscapes: A Case Study from Two Canadian Provinces" *Land* 13, no. 1: 6. <https://doi.org/10.3390/land13010006>

R Core Team. R: A Language and Environment for Statistical Computing; R Core Team: Vienna, Austria, 2023.

Deblonde, C., Cocking, R.B., Kerr, D.E., Campbell, J.E., Eagles, S., Everett, D., Huntley, D.H., Inglis, E., Parent, M., Plouffe, A., Robertson, L., Smith, I.R., and Weatherston, A., 2019. Surficial Data Model: the science language of the integrated Geological Survey of Canada data model for surficial geology maps; Geological Survey of Canada, Open File 8236, ver. 2.4.0, 1 .zip file. <https://doi.org/10.4095/315021>

Ridgeway, G. 2024. gbm: Generalized Boosted Regression Models. R Package, Version 2.1.9. <https://cran.r-project.org/web/packages/gbm/gbm.pdf>

R Core Team (2024). *R: A Language and Environment for Statistical Computing*.  
R Foundation for Statistical Computing, Vienna, Austria. <<https://www.R-project.org/>>.

Brandt, J.P. 2009. The extent of the North American boreal zone. *Environmental Reviews* 17: 101-161.

J.H. Friedman (2001). "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of Statistics* 29(5):1189-1232.