Bootleg Fire: Evaluating the Role of Fuel Treatments in Mitigating Burn Severity near Sycan Marsh, Oregon

Astrid Sanna & Jimmy Fowler

Table of content

GitHub repo
Check out our notebook!

Defining key terms: fuel treatments (fuel reduction)

Mechanical thinning (Tx)

- Reduces canopy density
- coarse and fine woody debris remain on the ground.

Prescribed fire (Rx)

- Reduces ground and surface fuels

Mechanical thinning + Prescribed fire (TxRx)

- You know it!

https://images.app.goo.gl/RZhYh6Fs8BMbAXKs6

Defining key terms: burn severity

Fire-suppressed forest

Ecologically-managed forest

no

Fuel treatments

yes

Crown-fire

Wildfire

Surface-fire

High

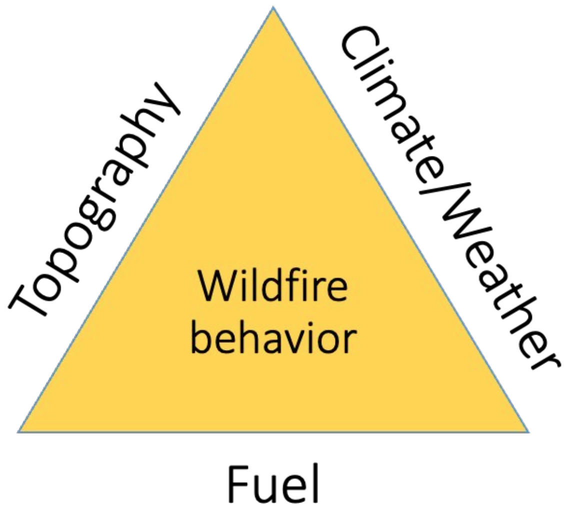
Burn-severity

Low/moderate

Relativized Burn Ratio (RBR)

https://www.nature.org/en-us/about-us/where-we-work/united-states/idaho/stories-in-idaho/wildfires-and-forest-management/

Burn severity by treatment type



Research question

Which environmental variables drove burn severity and what was the role of fuel treatments in mitigating burn severity?

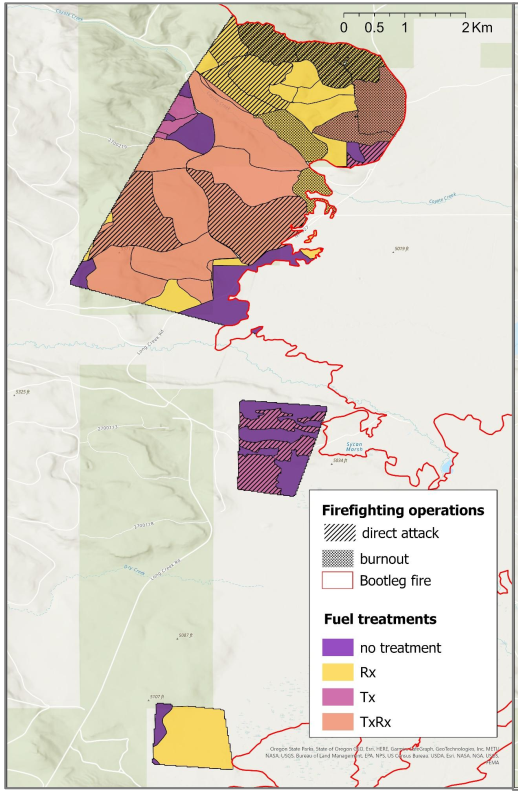
Objective

Quantifying the global and local importance of burn severity drivers.

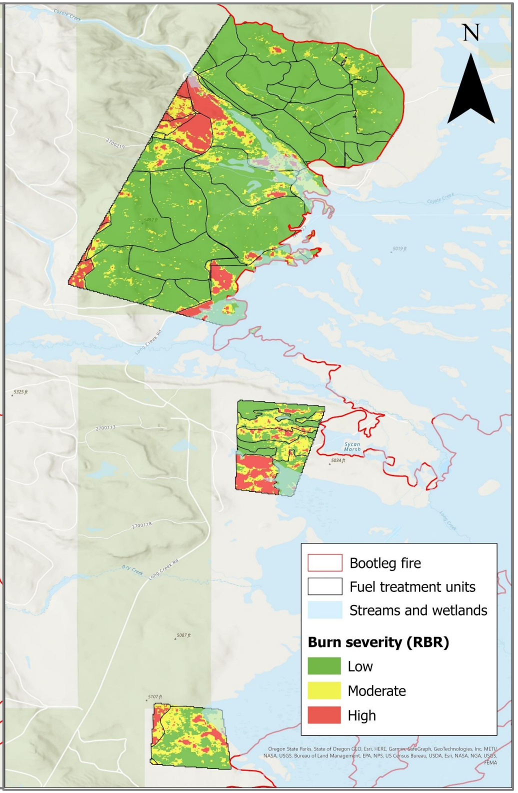
Methods: study area

- Dry-forest ecosystem
- Small section of the Bootleg fire, near Sycan Marsh, south-central OR
- Prescribed fire (Rx) and thinning + prescribed fire (TxRx) burned mostly at low severity with some patches of moderate/high severity.
- Thinning (Tx) and untreated units burned mostly at high severity with some patches of low/moderate severity.

Fuel treatments



Burn severity



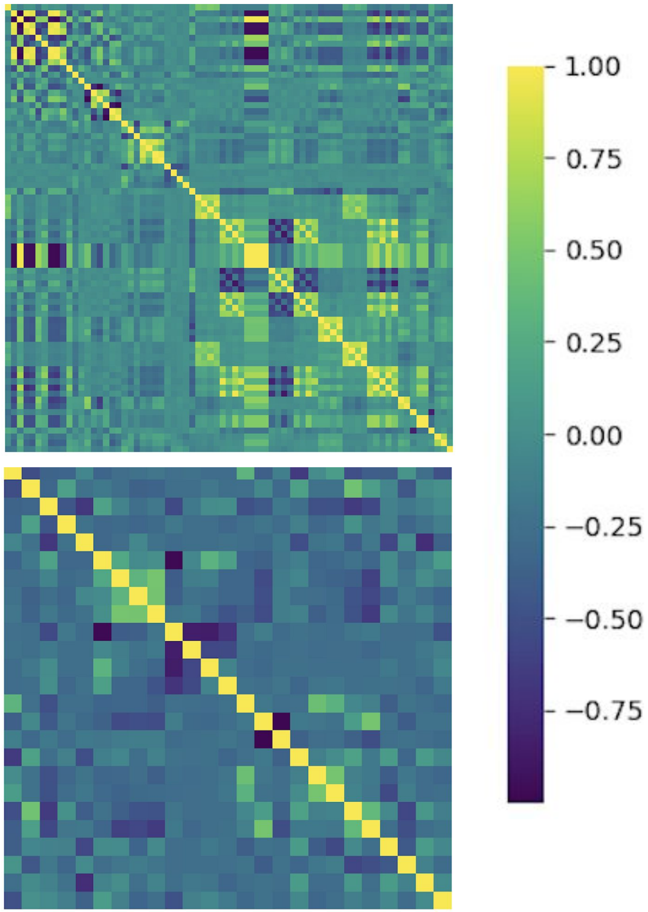
Methods: data frame building

- Variables (rasters):
  - Response: burn severity (RBR)
  - Predictors: topography, weather, forest structure, fuel treatments (presence/absence Rx)
- Sampling:
  - Created a point layer with points at 90-m distance (1825 points)
  - Used point layer to extract values from each variables (rasters)
- Data frame:
  - Combined all variable values in a single data frame
  - Retained points with canopy cover > 10%
  - Removed variables with correlation > 0.6

Correlation heatmaps

rho > 0.6
73
predictors

rho < 0.6
25
predictors

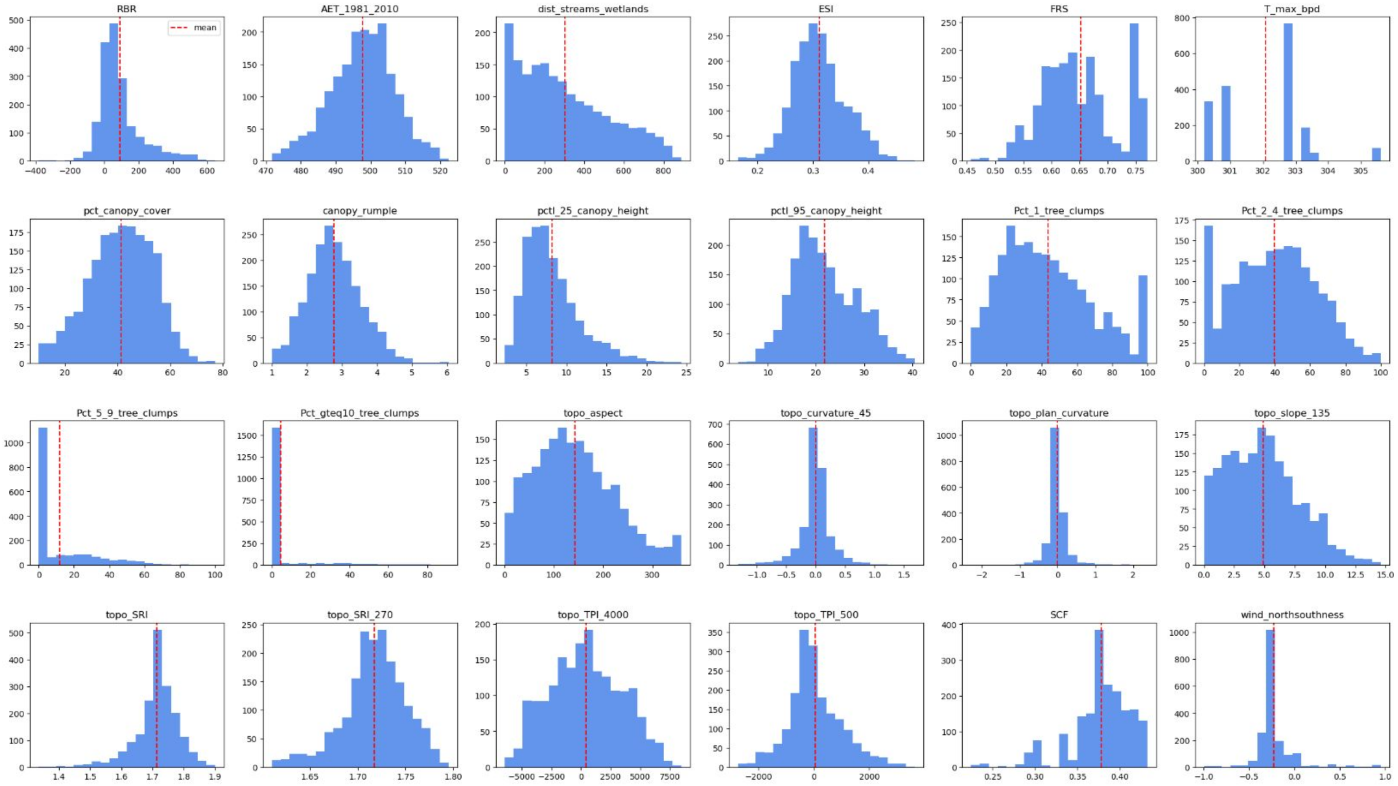


Methods: data characterization

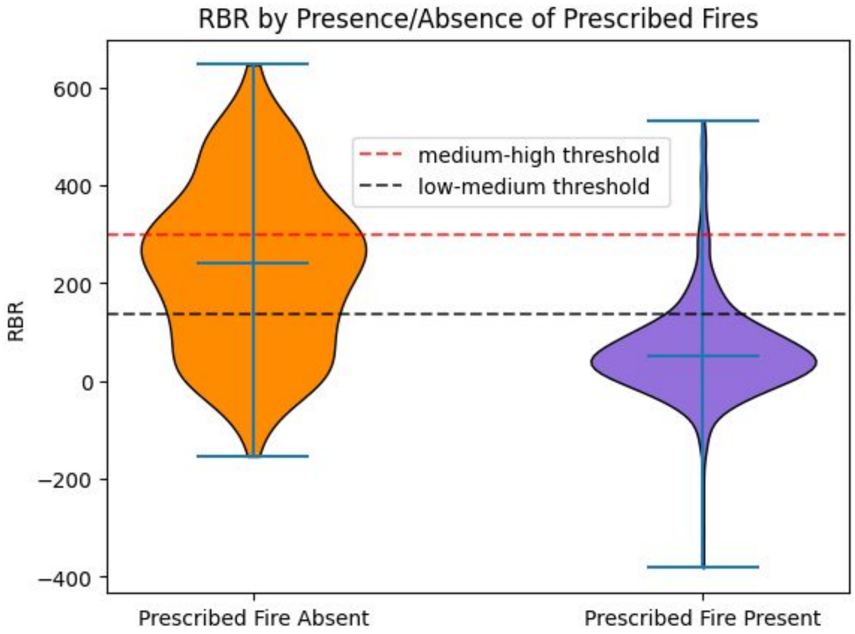
- Statistical Parameters
- Histograms
- Violin Plots
- Scatter Plots w/ LOWESS

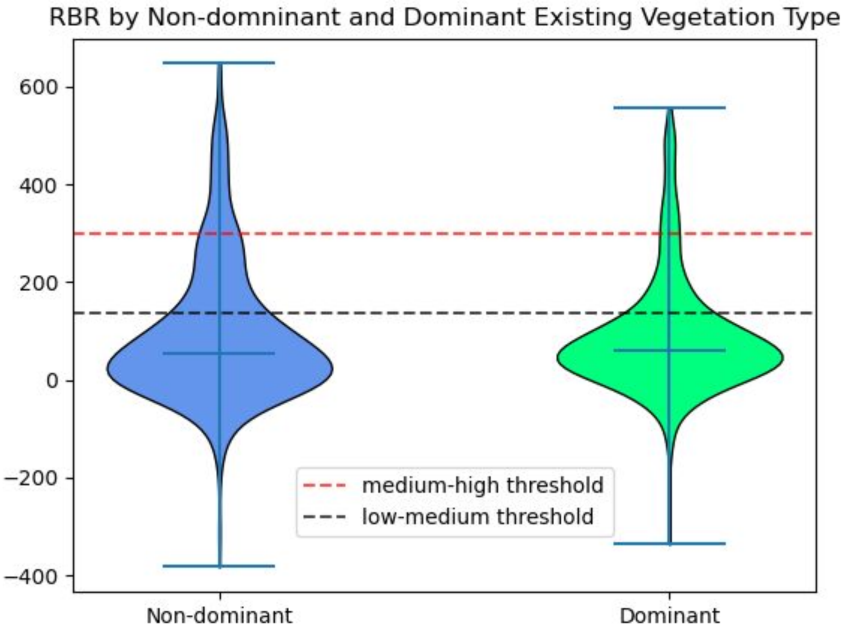
	mean	median	variance	kurtosis	skewness	Shapiro-Wilk
RBR	91.488219	60.000000	1.398217	2.479639	1.349423	1.970981e-35
AET_1981_2010	497.645108	498.058563	0.018374	-0.100091	-0.199147	6.841525e-06
dist_streams_wetlands	301.603877	256.512817	0.745351	-0.640629	0.588805	1.206388e-26
ESI	0.311384	0.307191	0.151182	0.080215	0.257402	1.279974e-08
FRS	0.652043	0.636571	0.101265	-0.748890	0.163477	1.261847e-21
T_max_bpd	302.073036	302.799988	0.004520	-0.584012	0.170595	1.033792e-38
pct_canopy_cover	41.327726	41.687901	0.292993	-0.403470	-0.203718	3.572863e-08
canopy_rumple	2.777495	2.719200	0.271477	0.189513	0.327035	9.251215e-08
pctl_25_canopy_height	8.192039	7.427400	0.422731	1.543436	1.168613	8.760923e-30
pctl_95_canopy_height	21.753912	20.788799	0.302518	-0.430169	0.339397	6.804063e-14
Pct_1_tree_clumps	43.648888	39.872741	0.580262	-0.444403	0.555046	7.627725e-23
Pct_2_4_tree_clumps	39.818124	40.338223	0.586714	-0.676311	0.077838	4.416008e-15
Pct_5_9_tree_clumps	11.895227	0.000000	1.542290	1.665613	1.544463	0.000000e+00
Pct_gteq10_tree_clumps	4.637761	0.000000	3.004746	10.861824	3.325000	0.000000e+00
topo_aspect	142.201424	133.407394	0.573051	-0.284807	0.490302	5.469322e-18
topo_curvature_45	0.003324	0.008000	82.601225	4.990927	-0.251907	5.397833e-33
topo_plan_curvature	-0.009161	-0.012700	-30.619895	17.385533	0.466060	7.006492e-45
topo_slope_135	4.874782	4.668100	0.605833	-0.262007	0.494386	1.932460e-18
topo_SRI	1.712883	1.716000	0.040334	2.807509	-1.021772	2.202023e-26
topo_SRI_270	1.717470	1.718900	0.018990	0.591671	-0.553570	2.575132e-16
topo_TPI_4000	406.527123	350.000000	7.468594	-0.723475	0.077223	2.811100e-11
topo_TPI_500	51.853699	-90.000000	18.779290	0.649399	0.411487	4.754146e-16
SCF	0.377831	0.380822	0.092749	2.284937	-1.224337	5.614855e-31
wind_northsouthness	-0.227879	-0.258820	-0.897839	11.767892	2.419382	0.000000e+00

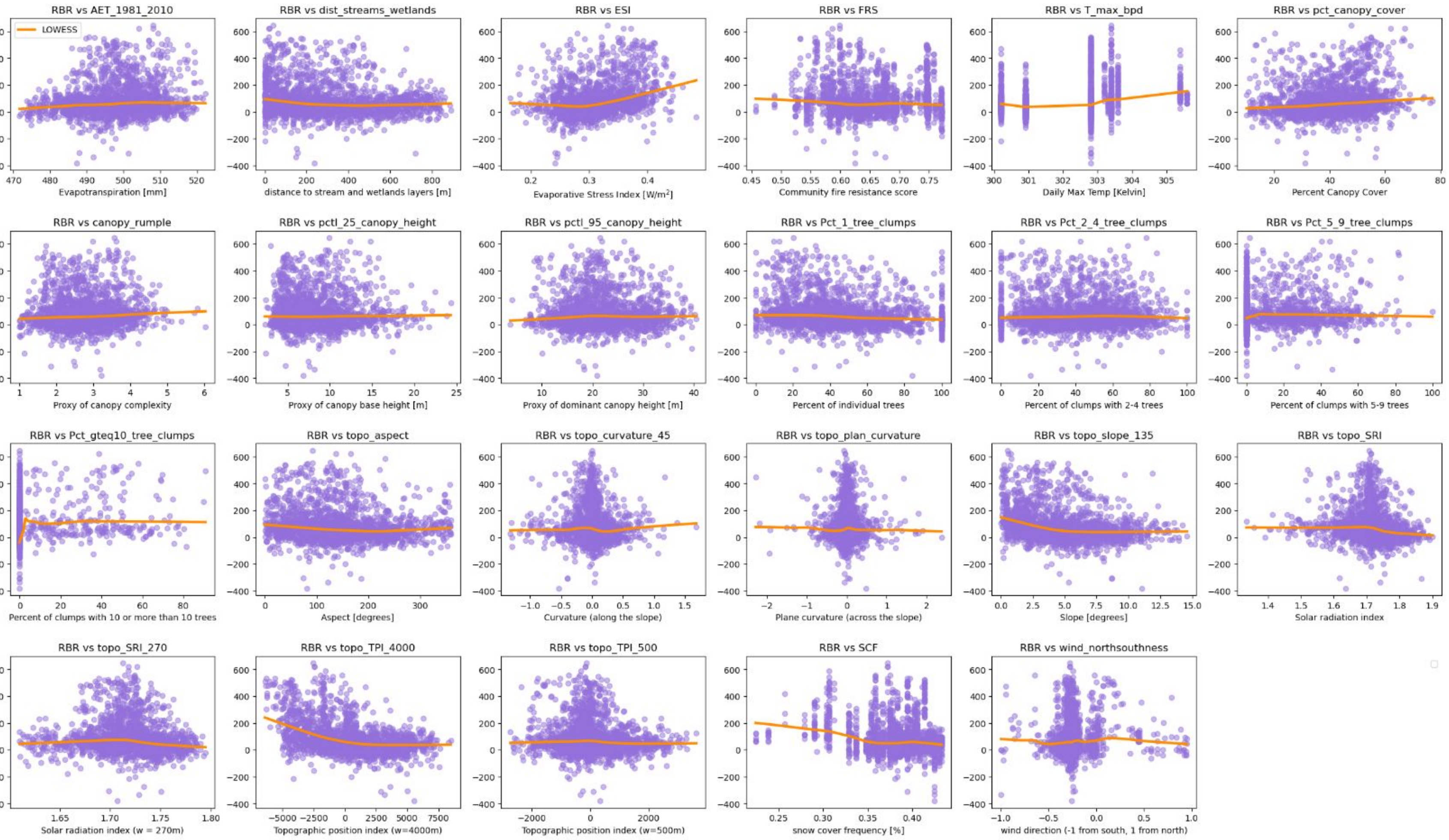
Methods: data characterization



Methods: data characterization







LOWESS
RBR vs AET_1981_2010
Evapotranspiration [mm]
RBR vs dist_streams_wetlands
distance to stream and wetlands layers [m]
RBR vs ESI
Evaporative Stress Index [W/m2]
RBR vs FRS
Community fire resistance score
RBR vs T_max_bpd
Daily Max Temp [Kelvin]
RBR vs pct_canopy_cover
Percent Canopy Cover
RBR vs canopy_rumple
Proxy of canopy complexity
RBR vs pctl_25_canopy_height
Proxy of canopy base height [m]
RBR vs pctl_95_canopy_height
Proxy of dominant canopy height [m]
RBR vs Pct_1_tree_clumps
Percent of individual trees
RBR vs Pct_2_4_tree_clumps
Percent of clumps with 2-4 trees
RBR vs Pct_5_9_tree_clumps
Percent of clumps with 5-9 trees
RBR vs Pct_gteq10_tree_clumps
Percent of clumps with 10 or more than 10 trees
RBR vs topo_aspect
Aspect [degrees]
RBR vs topo_curvature_45
Curvature (along the slope)
RBR vs topo_plan_curvature
Plane curvature (across the slope)
RBR vs topo_slope_135
Slope [degrees]
RBR vs topo_SRI
Solar radiation index
RBR vs topo_SRI_270
Solar radiation index (w = 270m)
RBR vs topo_TPI_4000
Topographic position index (w=4000m)
RBR vs topo_TPI_500
Topographic position index (w=500m)
RBR vs SCF
snow cover frequency [%]
RBR vs wind_northsouthness
wind direction (-1 from south, 1 from north)

End of Data Characterization

- Data not normally distributed
- No linear relationship with response
- There are many variables

We should use Random Forests Regression!

Methods: Random Forests

sklearn.ensemble.RandomForestRegressor

class `sklearn.ensemble.`**`RandomForestRegressor`**(*n_estimators=100, *, criterion='squared_error', max_depth=None, min_samples_split=2, min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features=1.0, max_leaf_nodes=None, min_impurity_decrease=0.0, bootstrap=True, oob_score=False, n_jobs=None, random_state=None, verbose=0, warm_start=False, ccp_alpha=0.0, max_samples=None*) [source]

Outline

- Check for NaNs (None)
- Split Train and Test datasets
- Tune hyperparameters
- Eliminate features using Recursive Feature Elimination with Cross-Validation (RFECV)
- Fit Random Forests model
- Evaluate predictive capabilities w/ R^2 and RMSE

Methods: Random Forests

```python
# Prepare data for train and test splits

# Define predictor features (X) and target variable (y)
X = df.drop(columns=['RBR'])  # 'RBR' is the target variable
y = df['RBR']

# Burn severity classes for stratification
thresholds = [135, 301] # low, medium, high burn severity
labels = np.digitize(y, thresholds, right=True)

# Check labels
label_counts = np.bincount(labels)

# Print the counts for each label
for label, count in enumerate(label_counts):
    print(f"Label {label}: {count} occurrences")

# Label 0: 1401 occurrences
# Label 1: 274 occurrences
# Label 2: 150 occurrences

# Split data in train and test by burn severity classes (labels)
X_train, X_test, y_train, y_test = train_test_split(X, labels,
                                                    test_size=0.2,
                                                    random_state=42,
                                                    stratify=labels)
```

Split into Train and Test Data

Hyperparameter Tuning

Feature Elimination

Fitting the Model

Validation

Methods: Random Forests

```
# Step 1: Hyperparameter Tuning (Number of Trees, Ma
param_grid = {
    'n_estimators': [200, 300, 500],  # Number of tr
    'max_depth': [10, 20, 30, 40, 'None'],  # Maximu
    'min_samples_leaf': [5, 10, 15, 20],  # Minimum
    'max_features': ['sqrt', 'log2', 'None'],  # Num
    'bootstrap': [True],
}

rf = RandomForestRegressor(random_state=42)

# Step 2: Hyperparameter tuning (GridSearchCV)
grid_search = GridSearchCV(rf, param_grid, cv=10,
                           scoring='r2', n_jobs=6)
grid_search.fit(X_train, y_train)

# Get the best hyperparameters
best_rf = grid_search.best_estimator_
```

Split into Train and Test Data

Hyperparameter Tuning

Feature Elimination

Fitting the Model

Validation

Methods: Random Forests

```
# Step 3: Feature selection using RFECV
rfecv = RFECV(estimator=best_rf, step=1, cv=10, scoring='r2')
rfecv.fit(X_train, y_train)

# Get the selected features
selected_features = X_train.columns[rfecv.support_]

# Subset your data to include only the selected features
X_train_selected = X_train[selected_features]
X_test_selected = X_test[selected_features]
```

Split into Train and Test Data

Hyperparameter Tuning

Feature Elimination

Fitting the Model

Validation

Methods: Random Forests

```
# Step 4: Fit the Random Forest Regression Model
final_rf = RandomForestRegressor(random_state=42, **grid_search.best_params_)
final_rf.fit(X_train_selected, y_train)
```

```
# Step 5: Evaluate the Model (R-squared and RMSE) on Train and
y_train_pred = final_rf.predict(X_train_selected)
y_test_pred = final_rf.predict(X_test_selected)

r2_train = r2_score(y_train, y_train_pred)
rmse_train = np.sqrt(mean_squared_error(y_train, y_train_pred))
r2_test = r2_score(y_test, y_test_pred)
rmse_test = np.sqrt(mean_squared_error(y_test, y_test_pred))

print(f"Selected Features: {selected_features}")
print("Best Hyperparameters:")
print(grid_search.best_params_)
print(f"R-squared (Train): {r2_train:.4f}")
print(f"RMSE (Train): {rmse_train:.4f}")
print(f"R-squared (Test): {r2_test:.4f}")
print(f"RMSE (Test): {rmse_test:.4f}")
```

Split into Train and Test Data

Hyperparameter Tuning

Feature Elimination

Fitting the Model

Validation

Methods: Random Forests - Model Performance

“None” <u>included</u> max_depth and max_features hyperparameters

R-squared (Train):	**0.8331**
RMSE (Train):	**0.2518**
R-squared (Test):	**0.5346**
RMSE (Test):	**0.4207**

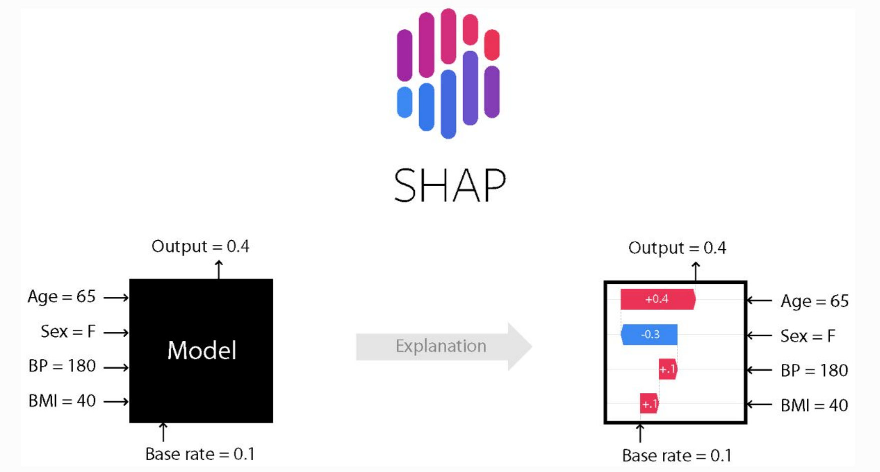
“None” <u>NOT included</u> in max_depth and max_features hyperparameters

Best Model

R-squared (Train):	**0.7901**
RMSE (Train):	**0.2824**
R-squared (Test):	**0.5436**
RMSE (Test):	**0.4166**

Methods: model explanation with SHAP analysis

- “**SHAP (SHapley Additive exPlanations)** is a game theoretic approach to explain the output of any machine learning model.”
- Offers local interpretability (importance of features for individual predictions) and global interpretability (aggregated importance across all predictions), while taking into account the interaction effects between predictors.



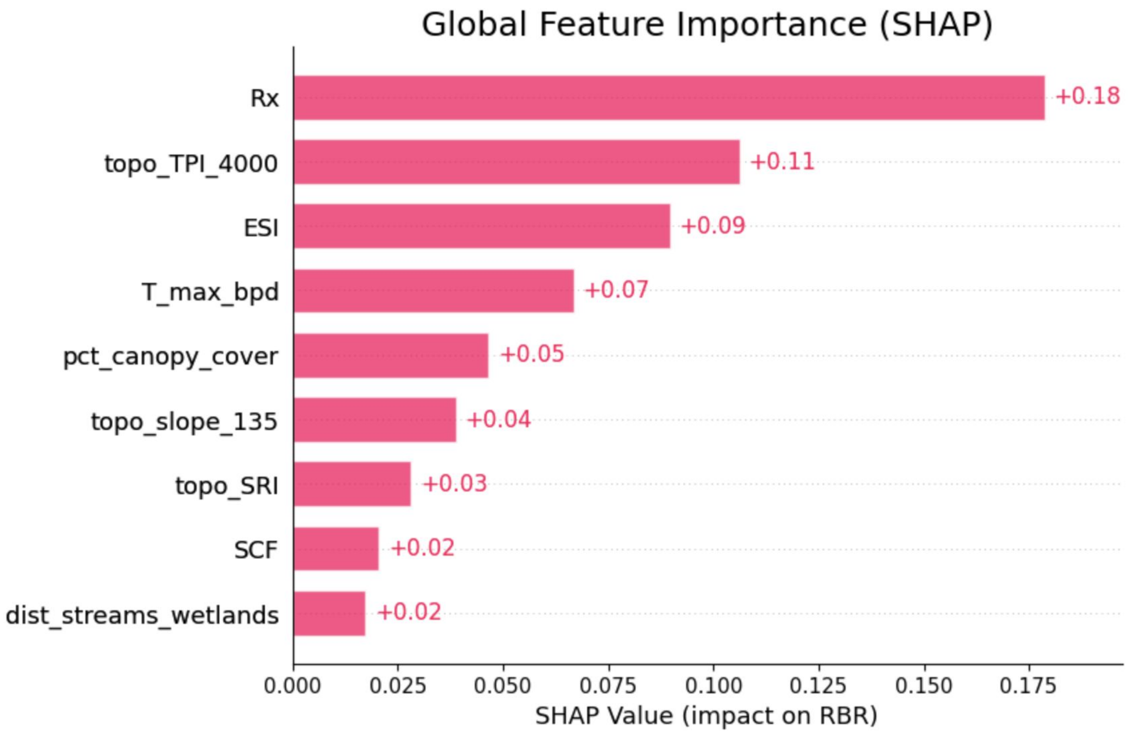
https://shap.readthedocs.io/en/latest/index.html

Results: RF regression model performance

- Retained method with best performance
- Performance sufficiently close to the expected R-squared (0.50-0.65)
- Moderate overfit to the train data, but not concerning since the model won’t be used for predictions.

Dataset	R-squared	RMSE
Train	0.79	0.28
Test	0.54	0.42

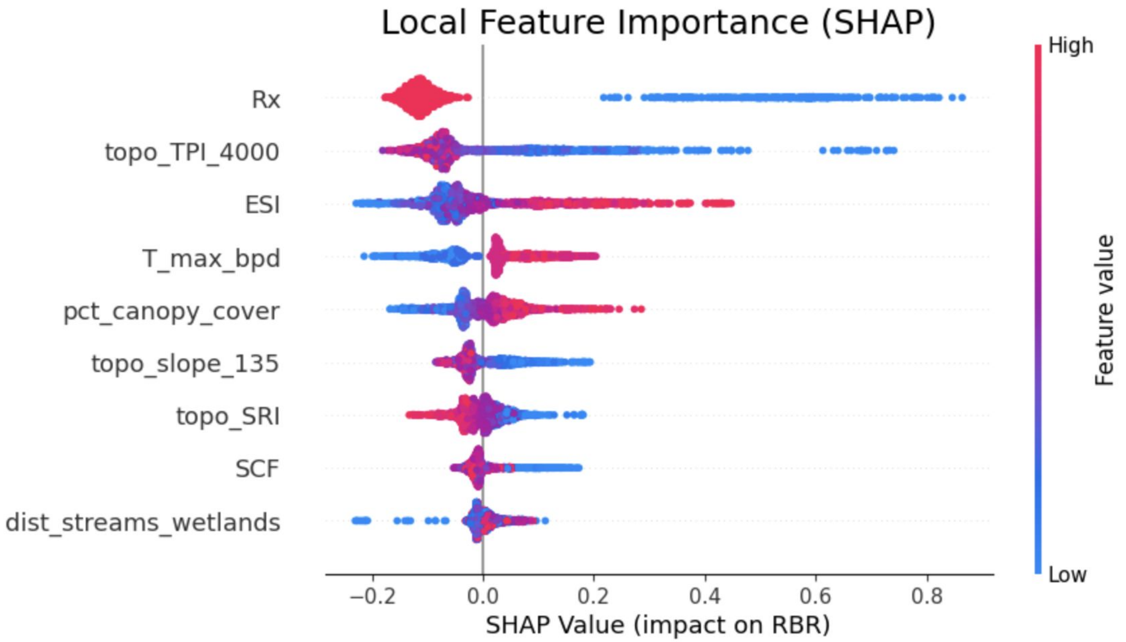
Results: global importance



First 5 key drivers of burn severity included:

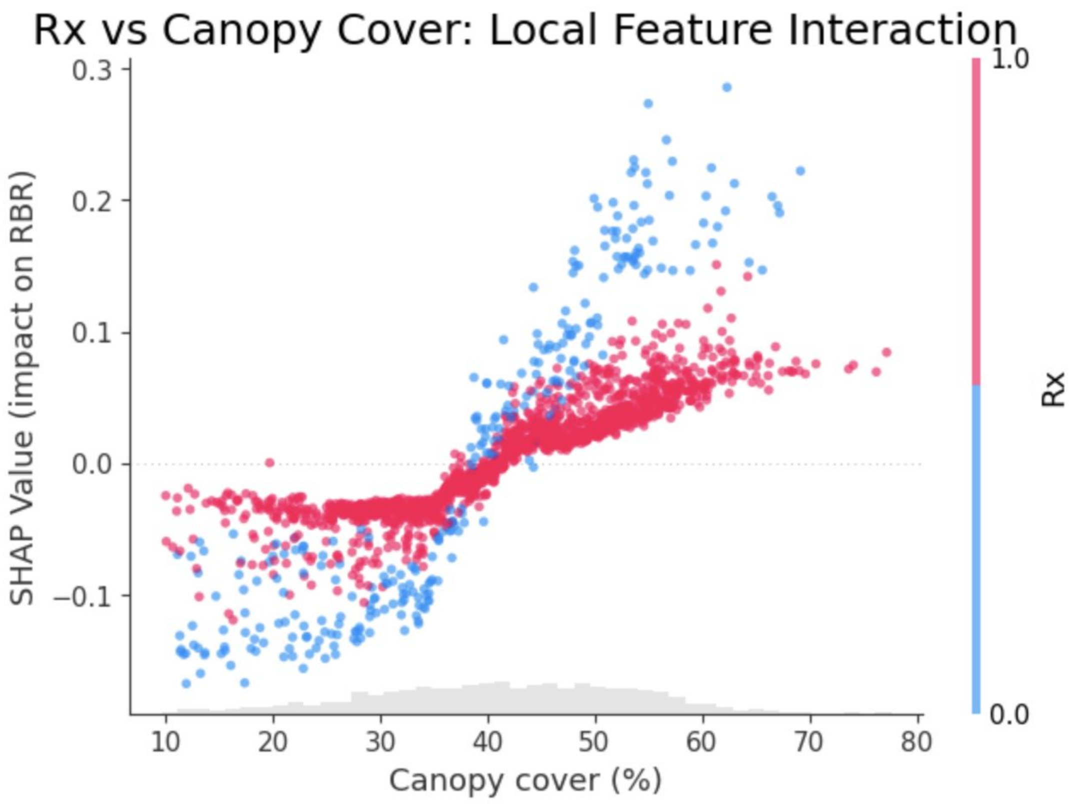
- prescribed fire
- coarse-scale topographic position index
- evaporative stress index
- max temperature by fire progression day
- percent canopy cover.

Results: local importance



- Drivers varied locally
- Where prescribed fire (Rx) was present, the model predicted lower fire severity.

Results: prescribed fire’s strongest interaction



- Prescribed fire (Rx) interacted more strongly with canopy cover
- At 40% and higher canopy cover, observations that received prescribed fire lead to a lower burn severity.

Lessons learned

- Preparing a dataset for a ML model is the most important and lengthy part – the ML code is just a couple of lines.
- SHAP analysis contributes to understanding and interpreting model performance, offering comprehensive insights (global and local perspective).
- Results suggested that the application of prescribed fire and its interaction with canopy cover mitigated burn severity by reducing ground, surface, and ladder fuel (preventing the fire to reach tree crowns).
- Fire weather and topography were also key drivers of burn severity and their interaction with prescribed fire must yet to be explored.

Study importance

The complex and dynamic nature of wildfires presents significant interpretive challenges. Identifying site-specific, key drivers of burn severity and the role of fuel treatments at mitigating burn severity are crucial tasks. Such insights guide the strategic allocation of limited management resources and lay the groundwork for future research.

The New York Times

July 19, 2021

How Bad Is the Bootleg Fire? It's Generating Its Own Weather.

THANK YOU!

Reach out if you have any questions astrid87@uw.edu!!