Predicting Post-wildfire Debris Flow Occurrence

Presented by: Thomas Oommen[1]

Priscilla Addison[1] , Ashley Kern[1] , Richard Coffman[2] , Sean Salazar[2], Thomas Oommen[1] , Keith Weber[3]

[1] Department of Geological and Mining Engineering and Sciences, Michigan Technological University, Houghton, MI

[2] Department of Civil Engineering, University of Arkansas at Fayetteville, Fayetteville, AR

[3] GIS Training and Research Center, Idaho State University, Pocatello, ID

Michigan Tech Create the Future

UNIVERSITY OF ARKANSAS

Idaho State UNIVERSITY

Background | I-5 30 mile (October 15, 2015)

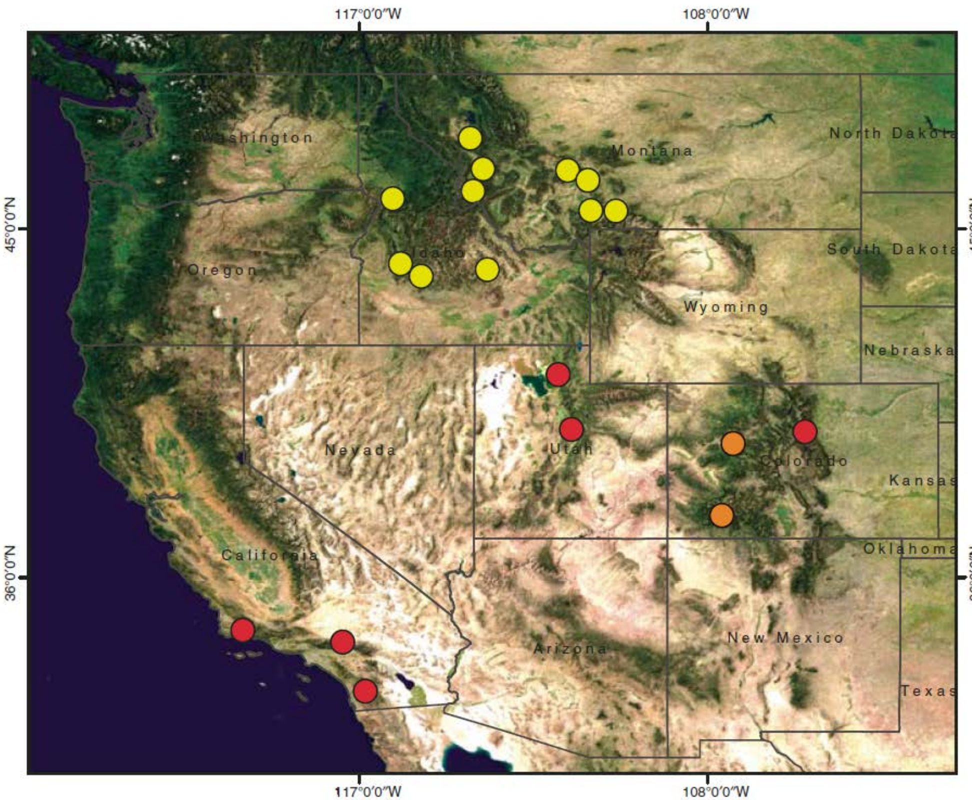


Figure 1. Map showing locations of basins used to develop models for the probability of debris-flow generation (yellow dots), for estimates of debris-flow volume (red dots), or both models (orange dots).

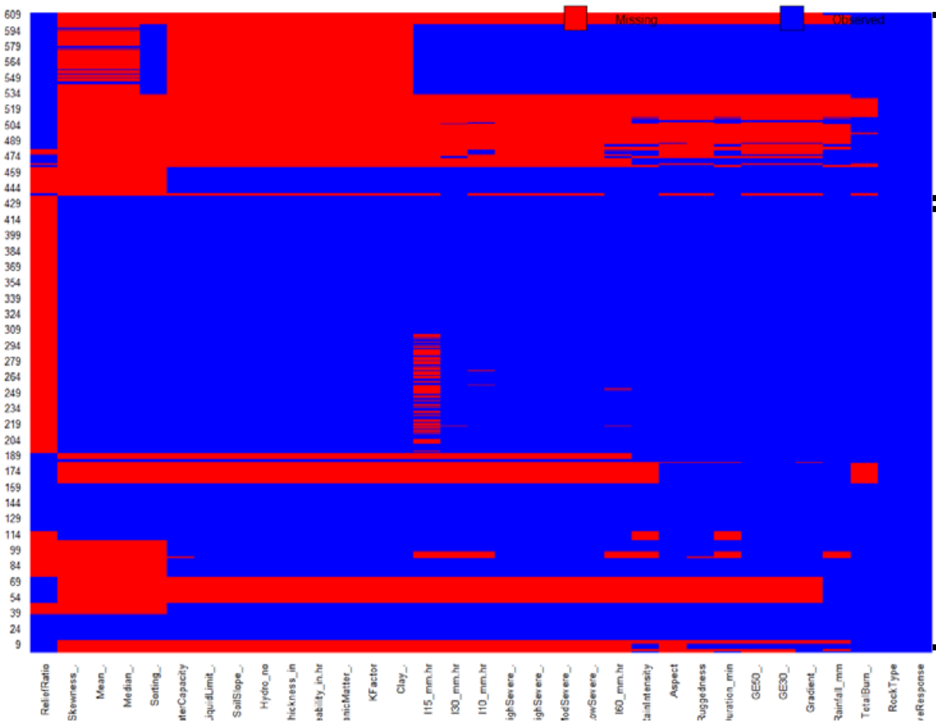
Intermountain

Cannon et al., 2010

California

Rupert et al., 2008 updated in 2011 by Susan Cannon

Background| Sparsity in Data



California *(144/66 samples)*

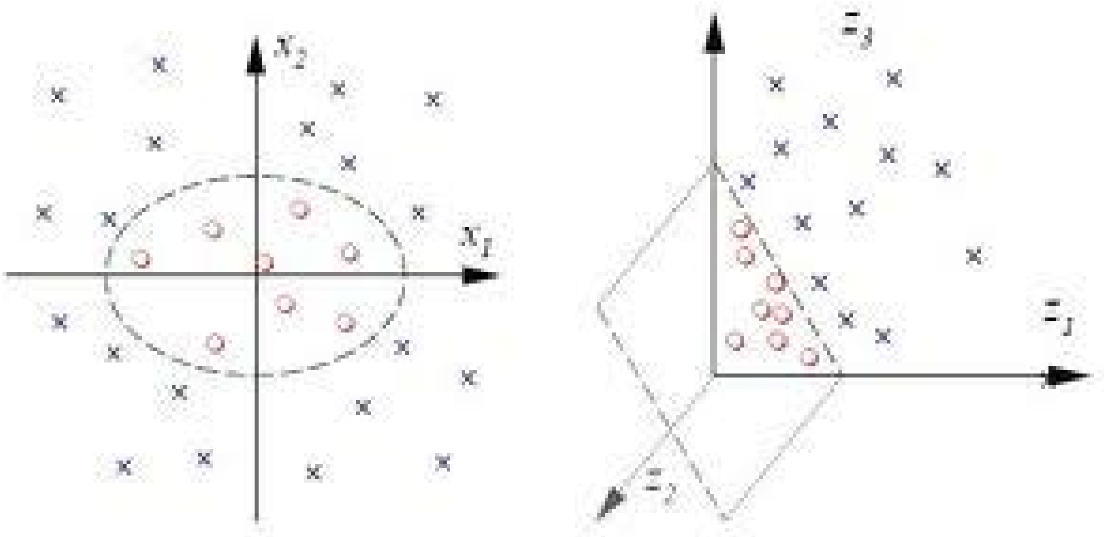
Intermountain West *(465/369 samples)*

33 predictors

609 sample points

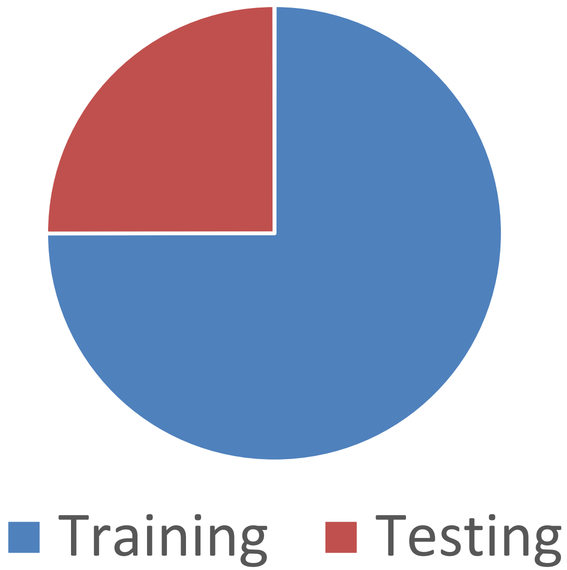
Background| Logistic Regression

- Basic approach that uses a logit function
- Advanced non-linear machine learning approaches
- Utilizing kernel functions

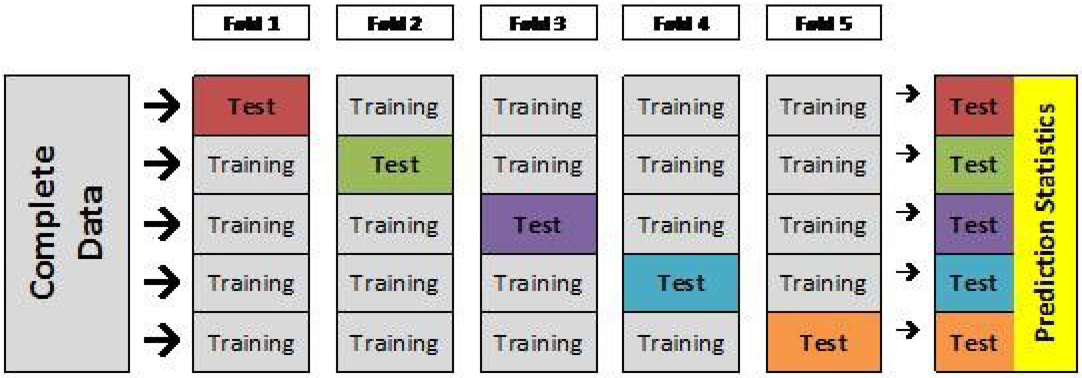


Background| Model Validation

- How well the developed model predicts?
- USGS models verify the predictive capability on the same data used for developing the model

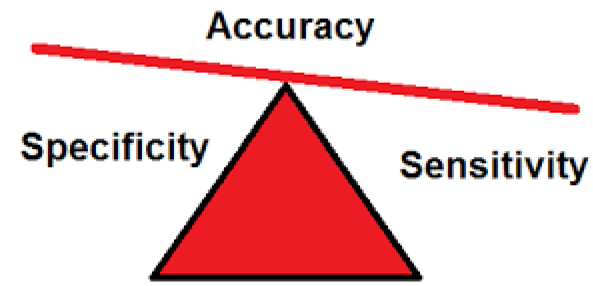


K-fold cross validation

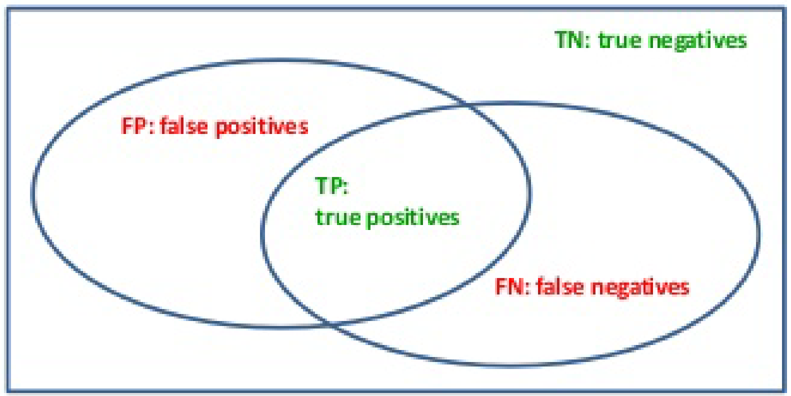


Background| Model Validation Statistics

	Debris flow	No-flow
Debris flow	85	5
No-flow	5	5



Overall accuracy = 85+5/(100) = 90%



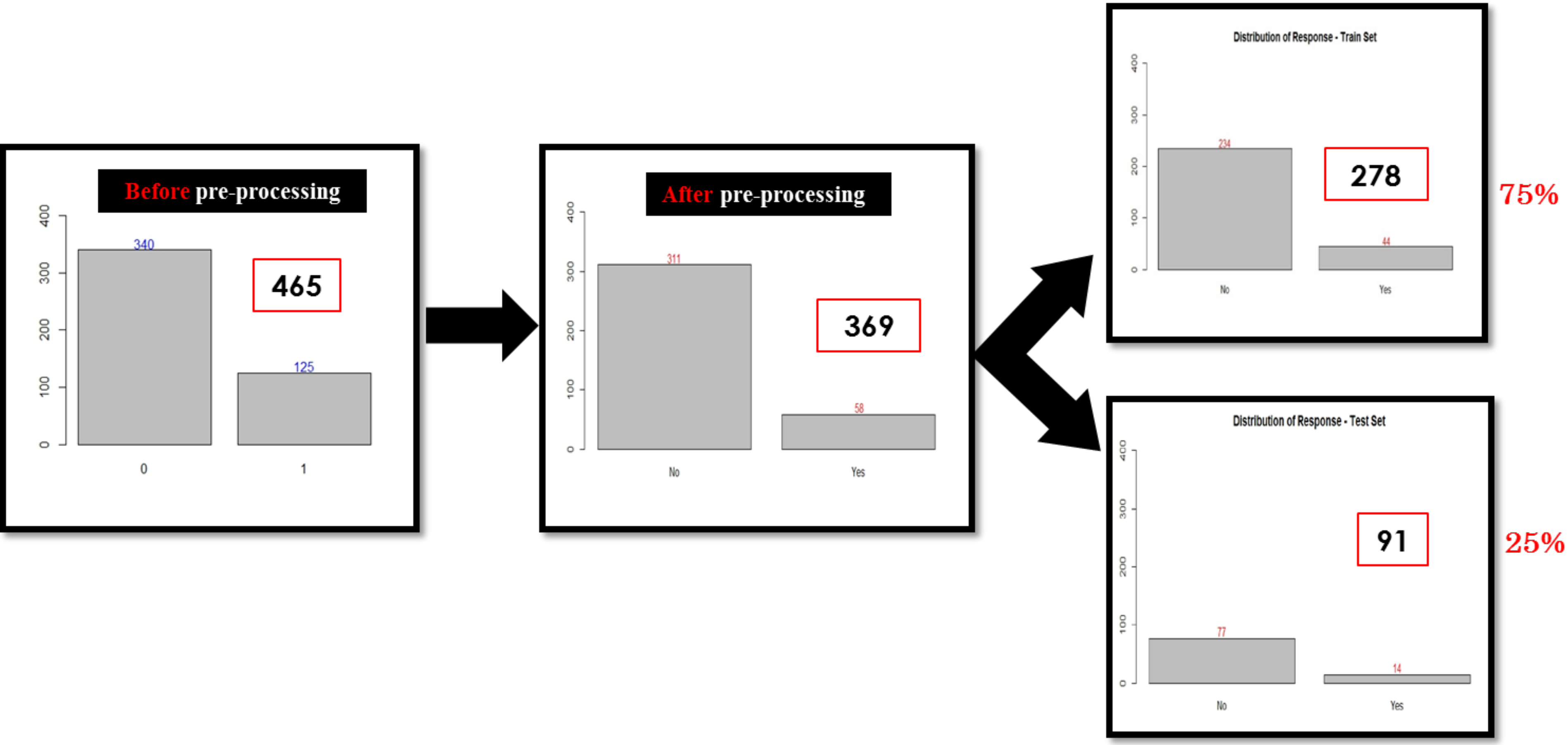
$$Sensitivity = \frac{TP}{TP + FN}$$

[same as recall;
aka **true positive rate**]

$$Specificity = \frac{TN}{TN + FP}$$

[aka **true negative rate**]

New Model| Intermountain



New Model| Intermountain

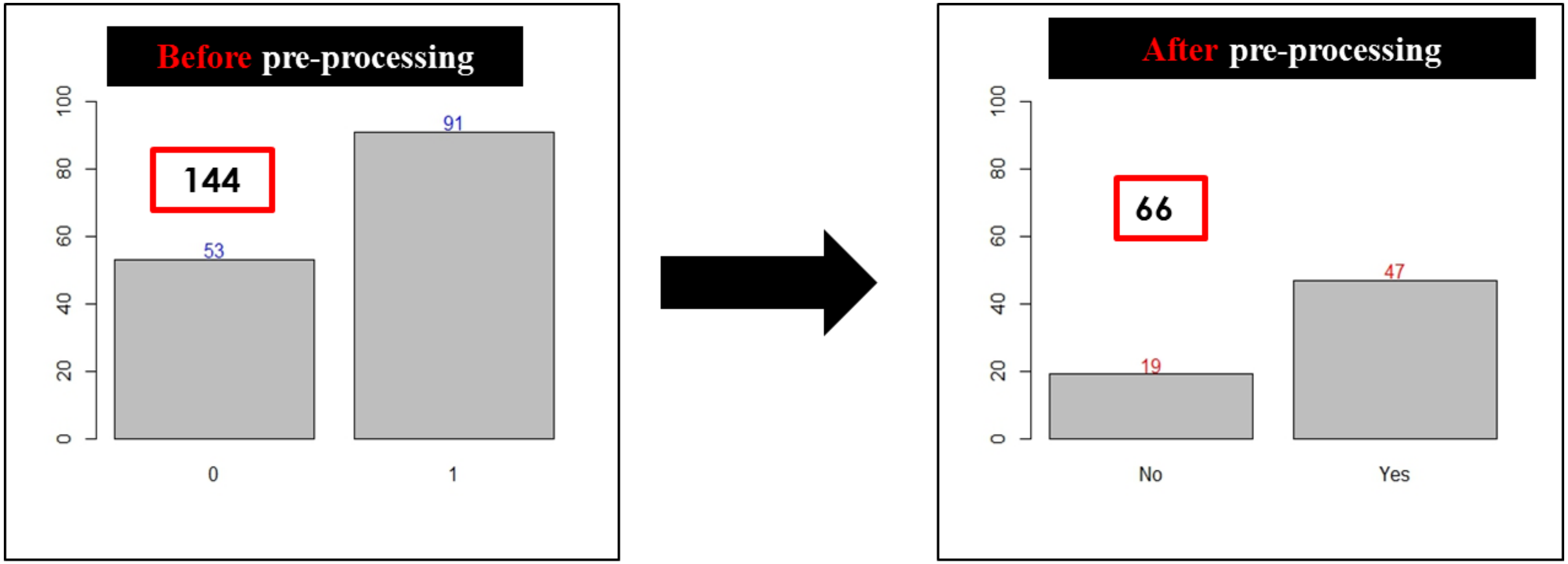
Validation Metrics (10 model runs)

MODEL	ACC.	SENS.	SPEC.
Logistic Regression (GLM)	0.86	0.42	0.94
Classification Trees Analysis (CTA)	0.85	0.46	0.92
Naïve Bayes (NB)	0.83	0.72	0.85
Mixture Discriminant Analysis (MDA)	0.86	0.71	0.89

Kern A. N., Addison P., Oommen T., Salazar S. E., & Coffman R. A., (2016) Machine learning based predictive modeling of debris flow probability following wildfire in the Intermountain Western United States. Mathematical Geosciences (Accepted for publication).

GLM = Cannon et al., 2010

New Model| California



10 fold cross validation repeated 10 times

New Model| California

Resampled Metrics – 10 fold CV

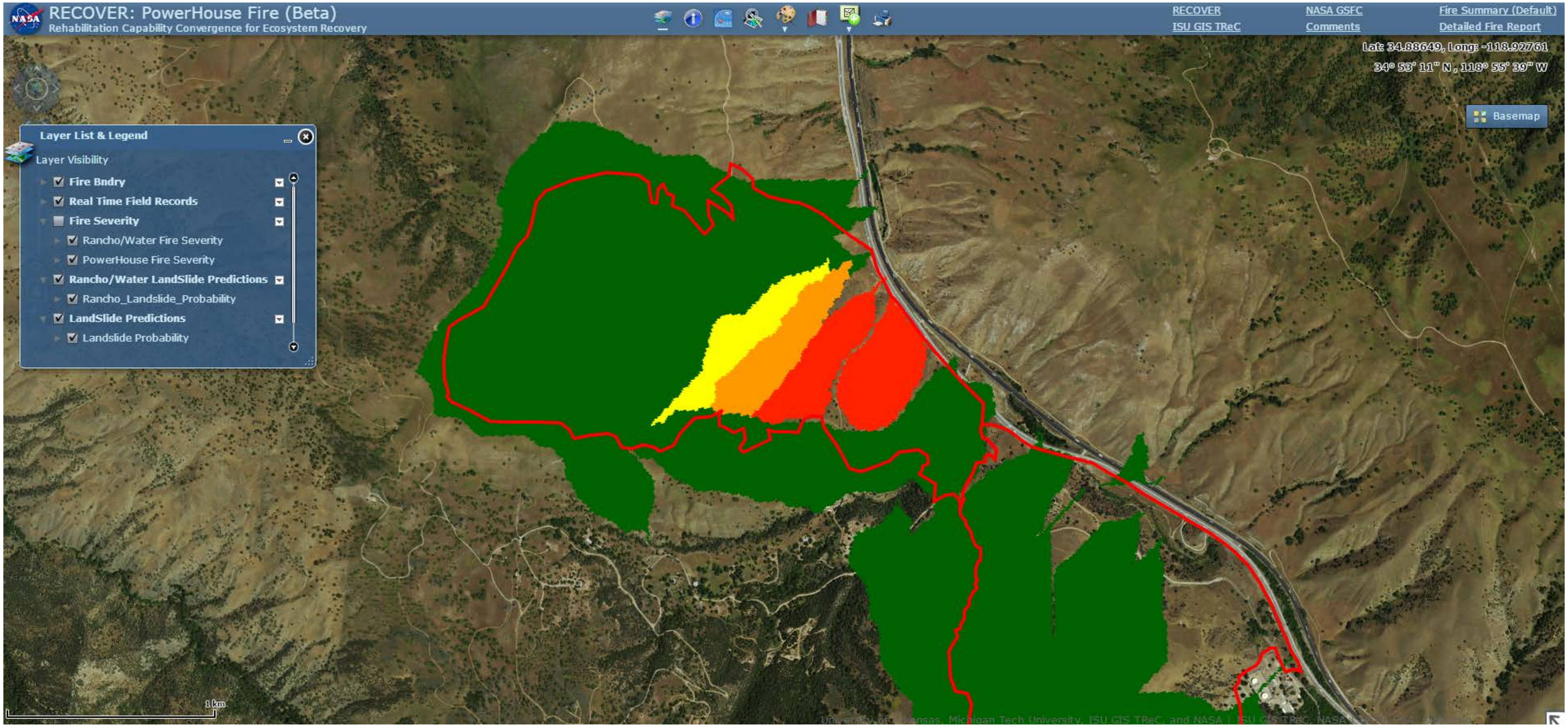
MODEL	ACC	SENS.	SPEC.
Logistic Regression **(GLM)**	0.49	0.36	0.80
Linear Discriminant Analysis **(LDA)**	0.49	0.35	0.83
Naïve Bayes **(NB)**	0.41	0.22	0.88
Averaged Neural Network **(ANN)**	0.46	0.30	0.86

Data too scanty to glean a representative trend

Summary

- Nonlinear models performing better than linear models suggest an underlying nonlinear relationship between predictors and response variable.
- Intermountain data performs better with a sensitivity of 72% for nonlinear Naïve Bayes model in comparison with 44% of existing logistic regression model by USGS (Cannon et al, 2010).
- California data too scanty to glean a trend from it. We recommend using the intermountain model for California until new refined model can be developed.

Research| Decision Support System



Final output can be delivered in few hours from request if inputs are available

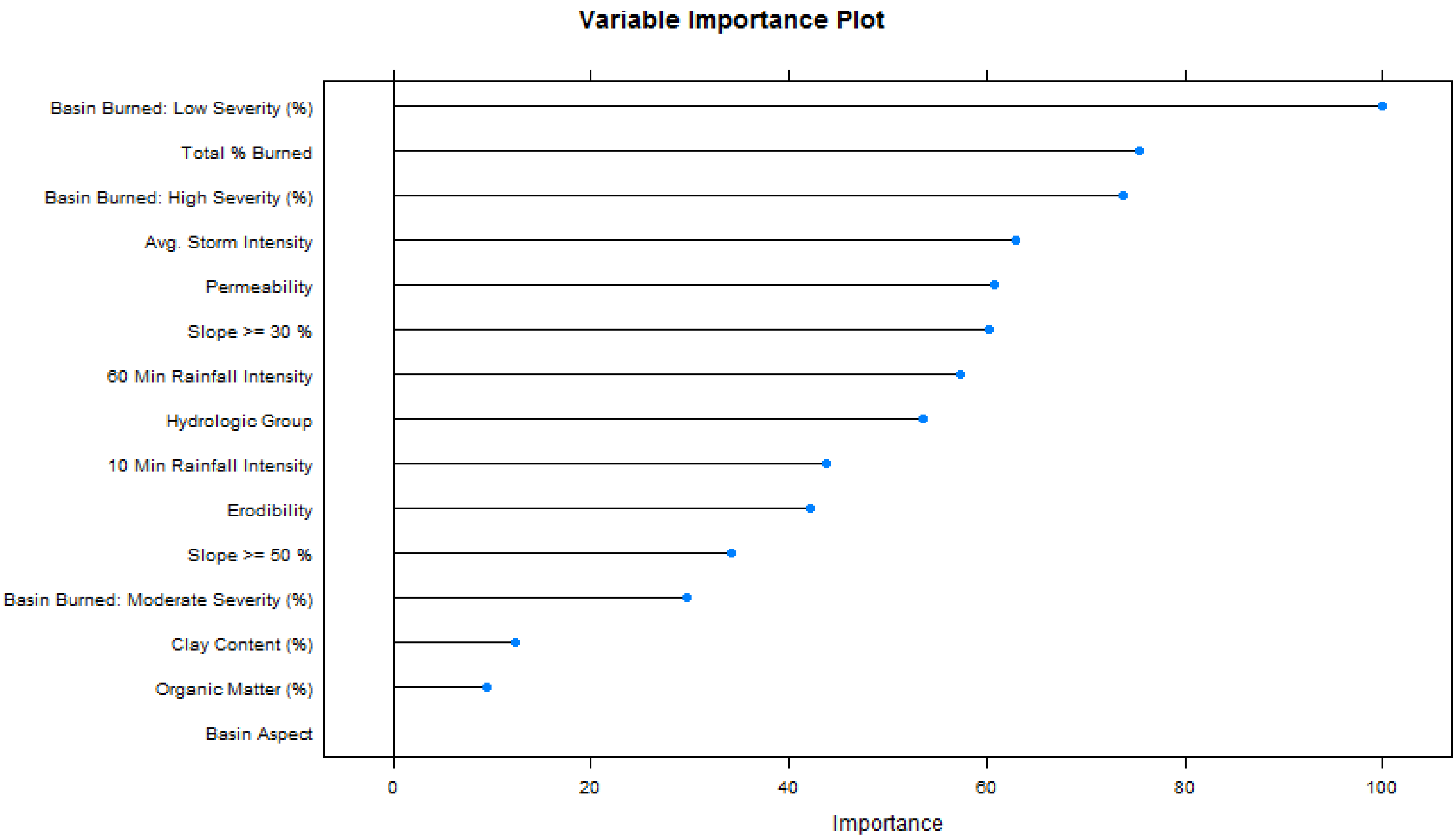
http://naip.giscenter.isu.edu/recover2/powerhousefire/

Acknowledgements

This research was made possible by the United States Department of Transportation (USDOT) *Office of the Assistant Secretary for Research and Technology (OST-R)* under Phase VI of the Commercial Remote Sensing & Spatial Information (CRS&SI) Technologies Program.

Variable Importance

Intermountain



Data Sources

No.	Variable	Description	Data	Source	Website
1	Basin burned: high severity (%)	Percent of the basin burned at high severity	Burn Severity	NASA RECOVER platform	http://giscenter.isu.edu/Research/Techpg/nasa_RECOVER/index.htm
2	Basin burned: moderate severity (%)	Percent of the basin burned at moderate severity			
3	Basin burned: low severity (%)	Percent of the basin burned at low severity			
4	Total basin burned (%)	Total percent of basin that has been burned			
5	Slope >= 30%	Burned basin area with slope >= 30 (%)	10m DEM	NRCS	https://gdg.sc.egov.usda.gov/
6	Slope >= 50%	Burned basin area with slope >= 50 (%)			
7	Basin Aspect	The average direction in degrees that the basin faces from north			
8	Hydrologic group	Infiltration rate for bare ground on a scale from 1 to 4; 1 = high infiltration, 4 = very slow	SSURGO		
9	Erodibility (k-factor)	Relative index of ability for soil to transport in rainfall			
10	Organic matter (%)	Percent of organic content in soil			
11	Clay content (%)	Percent of clay which is less than 2mm.			
12	Permeability	The rate at which water may flow through saturated soil	STATSGO		
13	Average storm intensity (mm/hr)	Average intensity of a single storm	Precipitation Frequency	NOAA atlas	http://hdsc.nws.noaa.gov/hdsc/pfds/
14	60 min rainfall intensity (mm/hr)	The 60 minute interval with the highest rainfall intensity			
15	10 min rainfall intensity (mm/hr)	The 10 minute interval with the highest rainfall intensity			