FINAL TECHNICAL REPORT FOR

"APPLICATION OF PRINCIPLED SEARCH MECHANISMS IN CLIMATE INFLUENCES AND MECHANISMS"

(NCC 2-1399)

Performance Period April 1, 2003 to March 31, 2005

Clark Glymour

University of West Florida Institute for Human and Machine Cognition, Pensacola, FL

Submitted to:

Joseph C. Coughlan
NASA Ames Research Center
Mail Stop 242-4
Moffett Field CA 94035-1000

ihmc
INSTITUTE FOR HUMAN & MACHINE COGNITION

University of West Florida, 40 South Alcaniz Street, Pensacola, FL 32502

NASA NCC2-1399

Applications of Principled Search Methods in Climate Influences and Mechanisms

Statistical Forecasts of Wildfire

Introduction

Forest and grass fires cause economic losses in the billions of dollars in the U.S. alone [14]. In addition, boreal forests constitute a large carbon store; it has been estimated that, were no burning to occur, an additional 7 gigatons of carbon would be sequestered in boreal soils each century [10]. Effective wildfire suppression requires anticipation of locales and times for which wildfire is most probable, preferably with a two to four week forecast [13], so that limited resources can be efficiently deployed. The United States Forest Service (USFS), and other experts and agencies have developed several measures of fire risk combining physical principles and expert judgment [11, 12], and have used them in automated procedures for forecasting fire risk. Forecasting accuracies for some fire risk indices in combination with climate and other variables have been estimated for specific locations [1,8], with the value of fire risk index variables assessed by their statistical significance in regressions. In other cases, the MAPSS forecasts [23, 24] for example, forecasting accuracy has been estimated only by simulated data.

We describe alternative forecasting methods that predict fire probability by locale and time using statistical or machine learning procedures trained on historical data, and we give comparative assessments of their forecasting accuracy for one fire season year, April-October, 2003, for all U.S. Forest Service lands. Aside from providing an accuracy baseline for other forecasting methods, the results illustrate the interdependence between the statistical significance of prediction variables and the forecasting method used.

Data

The Terrestrial Observation and Prediction System (TOPS) [5,6] provides measures of the following variables for years 2000-2003 for the lower 48 states gridded to 8km (approximately 15,814 acres) resolution using a Lambert equal-area projection (Figure 1 and Figure 2):

FPAR (Fraction of Photosynthetically Active Radiation absorbed by vegetation)
LAI (Leaf Area Index)
TMIN (Minimum temperature over past 24 hours)
TMAX (Maximum temperature over past 24 hours)
PRECIP (Amount of precipitation, rain or snow, over past 24 hours)
VPD (Vapor Pressure Deficit; an inverse function of humidity)

FPAR and LAI measures are collected every 8 days from NASA MODIS satellites [15]. The remaining variables are produced daily from ground observations collected by the National Climate Data Center (NCDC).

Fire forecasting models have been designed to predict number of fire days [1], or probability of at least one fire [8], or burned area [7], or fire "risk," the last a quantity that is not independently measurable and hence not amenable to an assessment of forecast accuracy. We estimate the probability of at least one fire in any 30 day period in the fire season within a 64 square kilometer grid cell. The National Interagency Fire Management Integrated Database (NIFMID) provides records of fires occurring on USFS land that required suppressive action for the years 1986-2003 [4], including fire location, ignition date and final fire size. In nearly all cases (99.9%), a fire lies entirely within the boundary of a single grid cell.

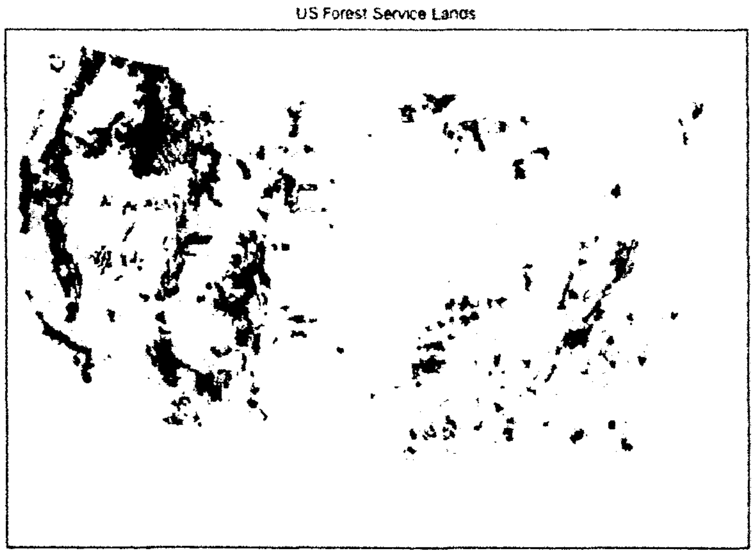


Figure 1: U.S. Forest Service Land

The USFS currently uses a map of fuel types (See Figure 2 and Table 1) in their wildfire assessment system [3,16] gridded to 1 km resolution.

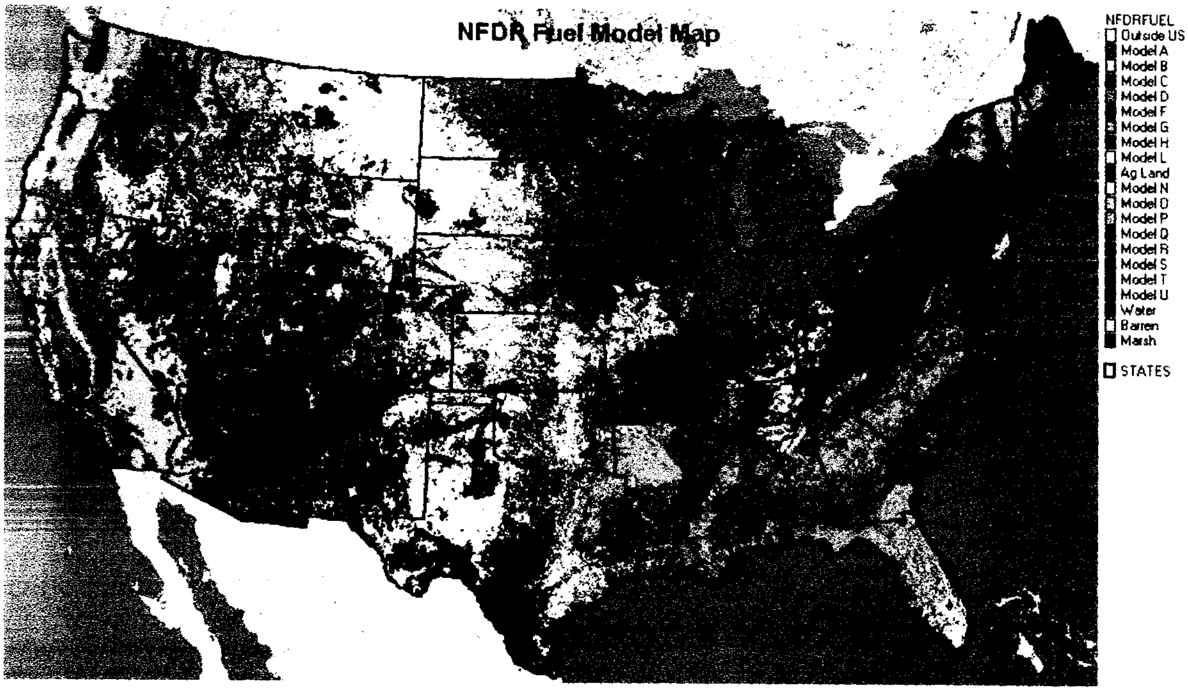


Figure 2

Code	#Cells	#USFS Cells	Description
A	615	21	Western Annual Grass
B	547	166	California Mixed Chaparral
C	7462	2938	Pine Grass Savanna
D	2024	105	Southern Rough
F	4370	944	Intermediate Brush
G	7805	4164	Short Needle Conifers with Heavy Dead Load
H	2602	1934	Short Needle Conifers with Normal Dead Load
L	21456	1593	Western Perennial Grasses
M	27233	843	Agricultural Land
N	450	7	Sawgrass or Other Thick Stemmed Grasses
O	930	158	High Pocosin
P	5892	716	Southern Pine Plantation
Q	623	33	Alaskan Black Spruce
R	19459	2312	Hardwoods
S	163	110	Alpine Tundra
T	15155	1833	Sagebrush-Grass Mixture
U	365	2	Western Long-Needle Conifer
V	861	13	Water
W	2995	75	Barren
X	69	1	Marsh

Table 1

For each year Y in 2000-2003, and each day-of-year T in 1-365 (except 2000, which begins on day-of-year 65), and each 64 sq. km. grid cell, values for the following variables were assigned:

- FPAR on day T, year Y
- LAI on day T, year Y
- TMIN on day T, year Y
- TMAX on day T, year Y
- PRECIP on day T, year Y
- VPD on day T, year Y
- TMIN averaged over [T-1,T-7], year Y
- TMAX averaged over [T-1,T-7], year Y
- PRECIP averaged over [T-1,T-7], year Y
- VPD averaged over [T-1,T-7], year Y
- TMIN averaged over [T-1,T-30], year Y
- TMAX averaged over [T-1,T-30], year Y
- PRECIP averaged over [T-1,T-30], year Y
- VPD averaged over [T-1,T-30], year Y
- Number of fires in calendar year Y-1
- Number of fires in calendar years [Y-2, Y-15]

All variables were standardized to zero mean and unit variance.

A Theremal Anomaly MODIS product is available which, if used in place of U.S. Forest Service fire records, would allow forecasting for the entire United State. The Thermal Anomaly product is very recent, however, and would not provide a substitute for number of fires years in calendar years, Y-2, Y – 15, a variable that proved to be significant.

Classifiers and Training

Three algorithms were trained: logistic regression [9], classification and regression trees [25], and support vector machines [22] with a radial basis kernel. All programs were coded using MATLAB libraries REF. In each case, training was with the 2000–2002 measures above. Logistic regression parameters were obtained with the Matlab Statistics Toolbox implementation, which performs 100 iterations of conjugate gradient search. Classification and regression parameters were obtained with the Matlab Statistics Toolbox implementation using the Gini measure and ten-fold cross-validation over ROC curve areas was used to prune the tree. Support vector machine tuning used default settings for LIBSVM 2.8. Supervision targets were scored as 1 if fire occurred within the succeeding 30 days, and 0 otherwise. Separate models from each classifier were trained for each month (e.g., March, April, May, ...) and for each fuel code. It is useful to know if recent weather events are on average important to forecast accuracy, or, alternatively, if the burn history for a place and day of year are accurate for a succeeding year. Accordingly, we also produced models that used burn history variables only as inputs.

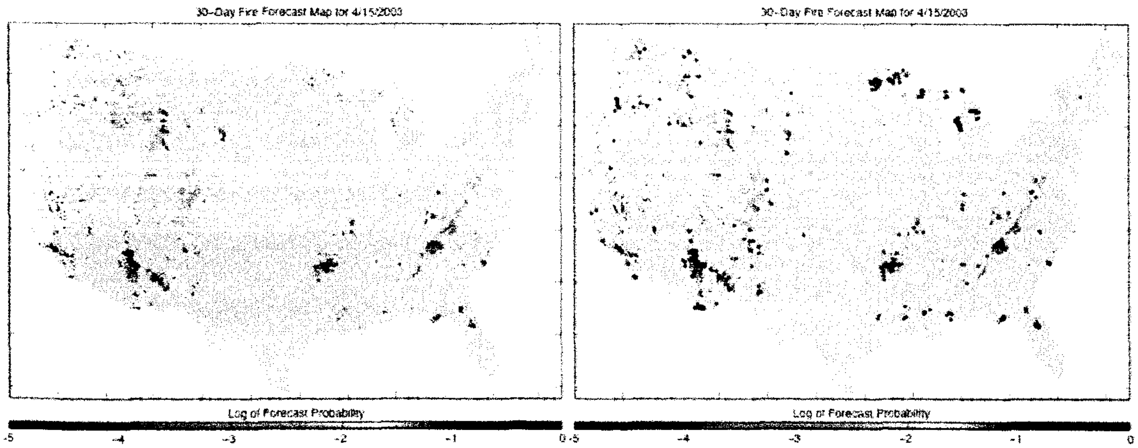
Testing Method

Each classifier was used to assign a fire occurrence probability to each grid cell in USFS lands for the 30-day period following the 15th day of each month from March through October of 2003. The NIFMID data for 2003 were then use to obtain ROC (Receiver Operating Characteristic) curves [18] for each fuel code, month and forecasting method. Areas under ROC curves (AUC) were calculated for each case. Maps of the resulting probabilities produced by one of the algorithms (logistic regression), and the actual fire occurrences, for April and July of 2003, are shown in Figure 3.

Results

The ROC curves for all forecasting methods, fuel codes and months for which at least 10 fires occur are shown in Appendix 1. Area under the curve (AUC) and number of acres burned are given in the legend for each figure. Tables 1-3 in Appendix 2 compare AUC values, fuel code by fuel code, with (top value) and without (bottom value) climate variables, for each of the 3 forecasting methods. Empty cells are for fuelcode months with fewer than 10 fires. Fuelcodes omitted altogether had fewer than 10 fires in any month. Cells where the AUC values differ significantly ($p < 0.05$ for a two-sided test) are highlighted in green, using Hanley's test [17, 18].

Tables 4-6 report all pairwise comparisons of the three classifier methods for each fuelcode month using both climate and burn history variables. Again, cells are left blank for fuelcode-months fewer than 10 fires. Cells are highlighted green if the first method significantly outperforms the second method, and they are highlighted magenta if the second method significantly outperforms the first. The same test and criterion for significance were used in these tables as those in Tables 1-3.



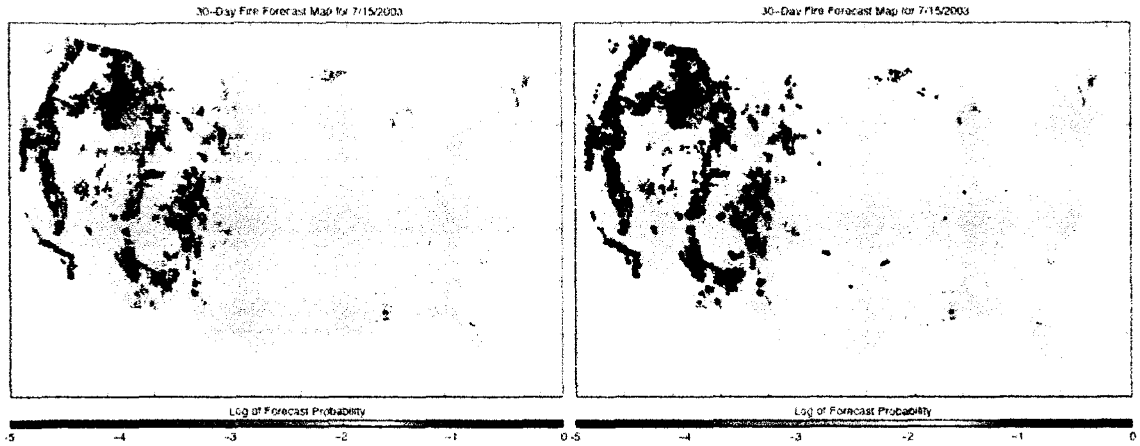


Figure 3

Discussion

The forecasting models tend to differ significantly in fuelcode-months with many fires. While climate variables are significant when support vector machine and CART classifiers are used, those variables are not significant when forecasting is with logistic regression, and on the whole, classification with logistic regression is substantially more accurate. Place is an effective proxy for many climate variables, and it may be that when associations of burn history and place are optimally exploited, on average climate variables are nearly independent of wildfire occurrence.

Several alternative explanations of our results are possible, but establishing any of them would require further extensive studies. The fire data we used in our study covers about 10% of the area of the lower 48 states, and different results might be obtained with fire data from a larger area; the MODIS Thermal Anomalies product is one possibility, although some problems with validation have been reported [19, 20]. Again, the value of climate variables for fire forecasts might be found in other studies, perhaps using other outcome criteria, such as acres burned or number of fire days rather than occurrence of at least one fire, but our results argue that such studies should consider multiple classification methods.

Except for precipitation, MODIS products that might serve in place of TOPS ground measured variables are presently available. Should the sufficiency of burn history for wildfire forecasting A MODIS burn product is recently available REF, which, over time, may permit global wildfire forecasting from burn history   Model training on a planetary scale is confounded by agricultural burnings, but these are typically either small fires or, when large, fires of known origin. With parallelization, and appropriate data, wildfire forecasts for all boreal forests, or for the entire Earth landmass, are computationally feasible.

References

[1] Andrews, P. L., D. O. Loftsgaarden, & L. S. Bradshaw. 2003. "Evaluation of Fire Danger Rating Indexes Using Logistic Regression and Percentile Analysis." *International Journal of Wildland Fire*, 12: 213-226.

[2] Bonnlander, B.V. 2005. "Statistical Forecasts of Wildfire: A Baseline Approach". 2005. CMU Laboratory for Symbolic Computation Technical Report #172.

[3] Burgan, R.E, Klaver, R.W., and Klaver, J.M. 1998. "Fuel Models and Fire Potential from Satellite and Surface Observations." *International Journal of Wildland Fire* 8(3):59-170.

[4] USDA Forest Service. 1993. National Interagency Fire Management Integrated Database (NIFMID) reference manual. U.S. Department of Agriculture, Forest Service, Fire and Aviation Management. Washington, D.C.-USA. 43 p.

[5] Thornton, P.E., Running, S.W., and White, M.A. 1997. "Generating Surfaces of Daily Meteorological Variables over Large Regions of Complex Terrain." *Journal of Hydrology* 190:214-251.

[6] White, M.A. and Nemani, R.R. 2004. "Soil Water Forecasting in the Continental United States: Relative Forcing by Meteorology Versus Leaf Area Index and the Effects of Meteorological Forecast Errors." *Canadian Journal of Remote Sensing*, 30(5):717-730.

[7] Medler, M.J. 1999. "Modeling Fire Hazards with Remotely Sensed Data from Recent Fires." *Proceedings of the Joint Fire Science Conference and Workshop*.

[8] Preisler, H.K., Brillinger D.R., Burgan R.E., Benoit J.W. 2004. "Probability Based Models for Estimation of Wildfire Risk." International Journal of Wildland Fire 13:133-142.

[9] Hosmer DW, Lemeshow S. *Applied logistic regression*. 2nd ed. New York: Wiley, 2000.

[10] Wardle D.A., Hornberg G., Zackrisson, M.K., Coomes, D.A. 2003. "Long-Term Effects of Wildfire on Ecosystem Properties Across an Island Area Gradient." *Science* 300:972-975.

[11] Westerling, A. L., A. Gershunov, D. R. Cayan, & T. P. Barnett. 2002. "Long Lead Statistical Forecasts of Area Burned in Western U.S. Wildfires by Ecosystem Province." *International Journal of Wildland Fire*, 11: 257-266.

[12] "Gaining an Understanding of the National Fire Danger Rating System". 2002. *NFES publication #2665*, http://www.nwcg.gov.

[13] Brown, T.J., G.G. Garfin, T. Wordell, R. Ochoa and B. Morehouse. 2004. "Climate, fuels, fire and decisions: The making of monthly and seasonal wildland fire outlooks." *American Meteorological Society 14th Conference on Applied Climatology, Proceedings*, 8 pp.

[14] Environmental Media Services website. http://www.ems.org/wildfires/costs.html.

[15] Overview of the MODIS Mission. PDF file at web address http://modis.gsfc.nasa.gov/about/media/modis_brochure.pdf.

[16] USDA Forest Service. 2005. WFAS: NFDRS Fuel Model Map. [Online] http://www.fs.fed.us/land/wfas/nfdr_map.htm [April 2005].

[17] Hanley JA & McNeil BJ. 1983. "A method of comparing the areas under Receiver Operating Characteristic curves derived from the same cases." *Radiology* 148, 839-843

[18] Hanley JA, McNeil BJ. 1982. "The meaning and use of the area under the Receiver Operating Characteristic (ROC) curve." *Radiology* 143, 29-36

[19] Giglio, L., Descloitres, J., Justice, C. O., and Kaufman, Y., 2003, An enhanced contextual fire detection algorithm for MODIS. *Remote Sensing of Environment*, 47:1311-1318.

[20] Kaufman, Y., Ichoku, C., Giglio, L., Korontzi, S., Chu, D. A., Hao, W. M., Li, R.-R., and Justice, C. O., 2003, Fires and smoke observed from the Earth Observing System MODIS instrument: products, validation, and operational use. *International Journal of Remote Sensing*, 24:1765-1781.

[21] Golub, G.H. and Van Loan, C.F. 1996. *Matrix Computations.* Johns Hopkins University Press.

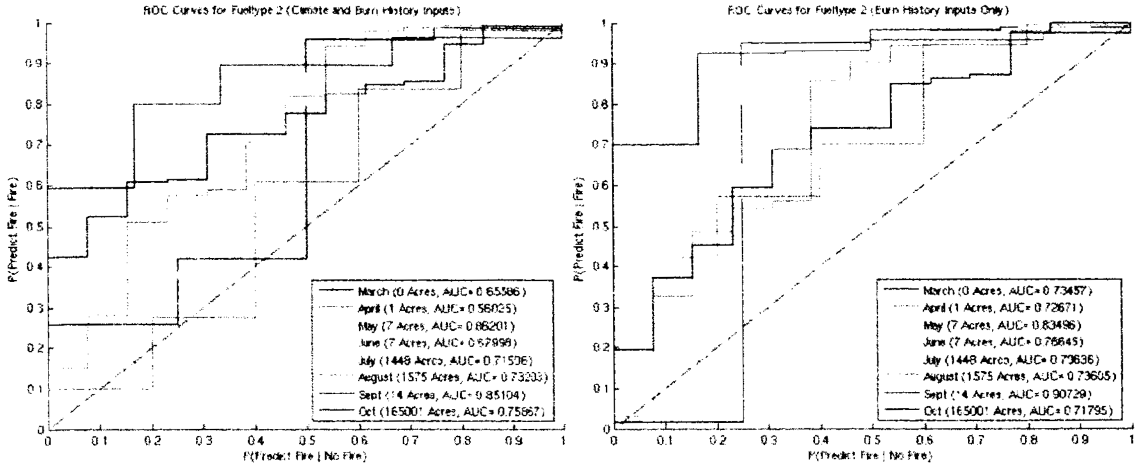
[22] Chang, C.C. and Lin, C.J. 2001 LIBSVM: a library for support vector machines. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[23] Lenihan, J.M., Drapek R.J., Bachelet D. and Neilson R.P. 2003. Climate change effects on vegetation distribution, carbon, and fire in California. *Ecological Applications* 13, 1667-1681.

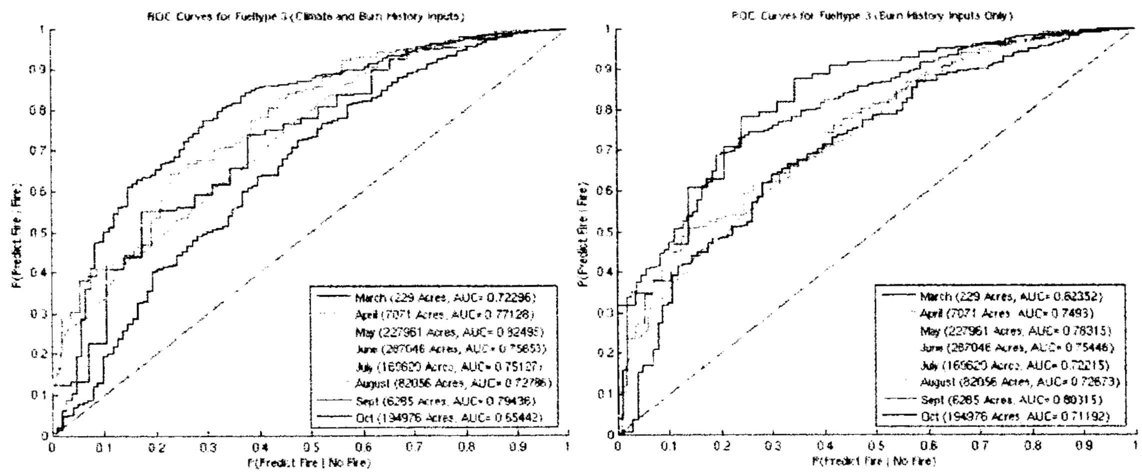
[24] Bachelet, D., J.M. Lenihan, C. Daly, R.P. Neilson, D.S. Ojima, W.J. Parton. 2000. MC1: A dynamic vegetation model for estimating the distribution of vegetation and associated ecosystem fluxes of carbon, nutrients, and water. *USDA General Technical Report PNW-GTR-508.*

[25] Breiman,. L., Friedman, J., Olshen, R. and Stone, C. 1984. *Classification and Regression Trees.* Chapman & Hall, NY.

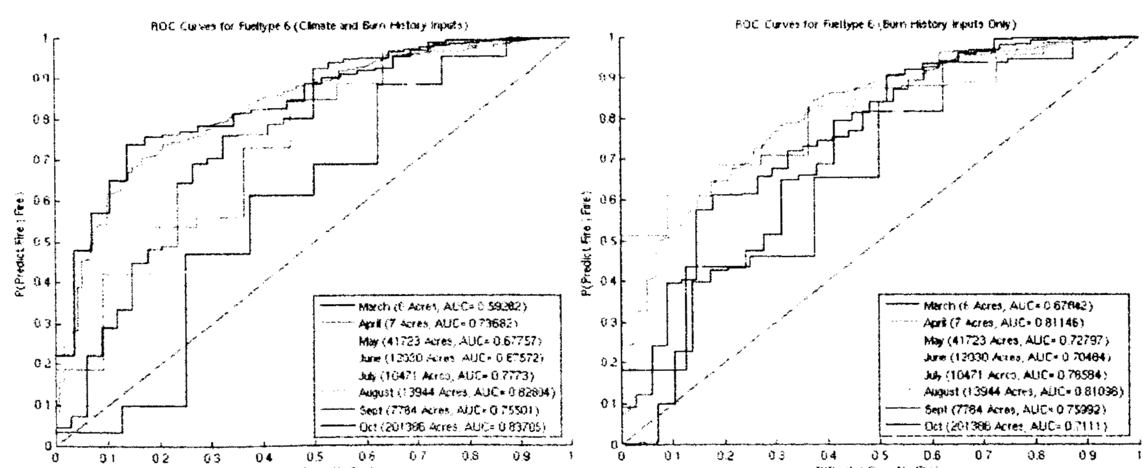
Appendix 1: ROC curves



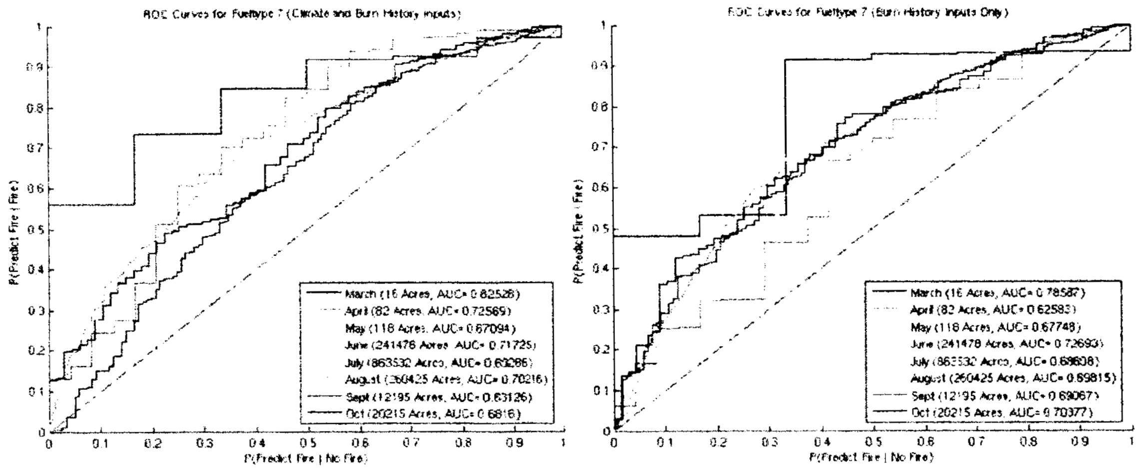
Fuel Code B: California Mixed Chaparral (logit method)



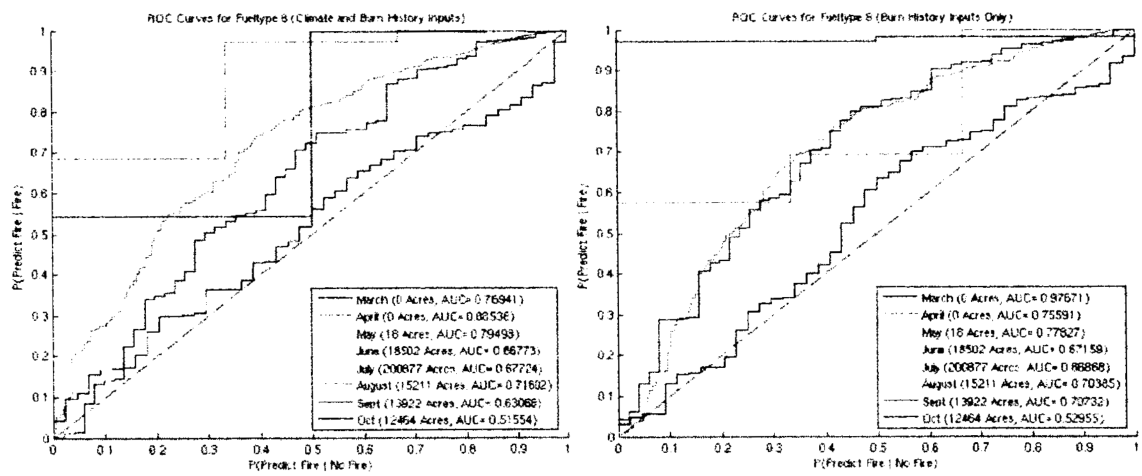
Fuel Code C: Pine Grass Savanna (logit method)



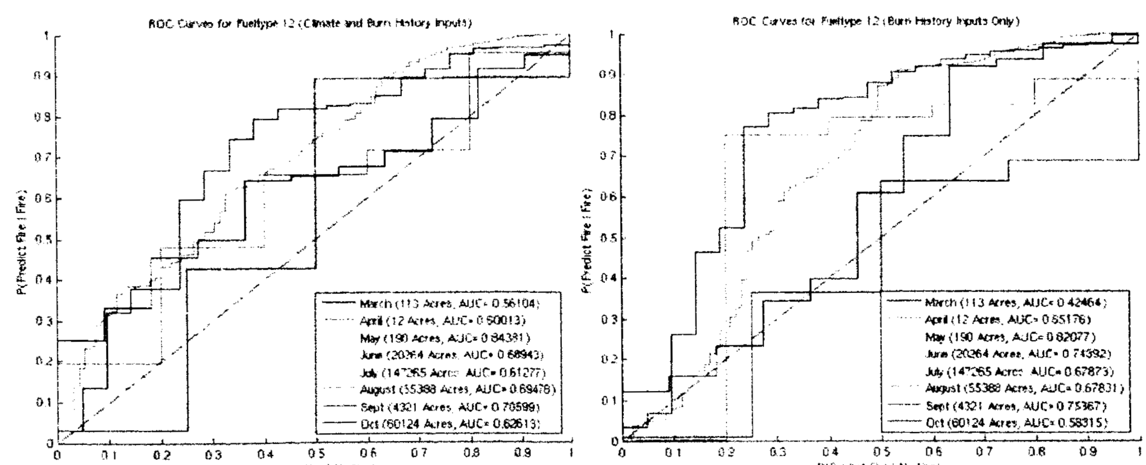
Fuel Code F: Intermediate Brush (logit method)



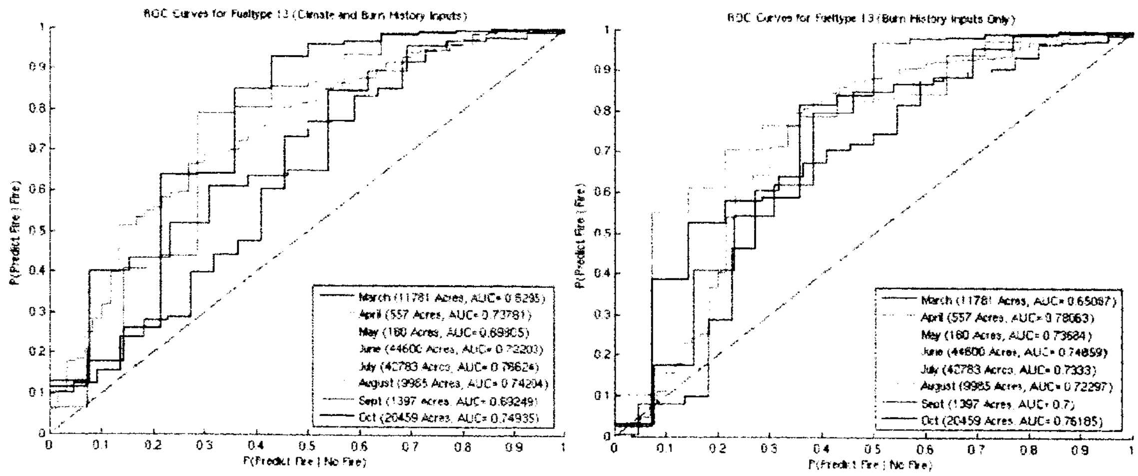
Fuel Code G: Short Needle Conifers, Heavy Dead Load (logit method)



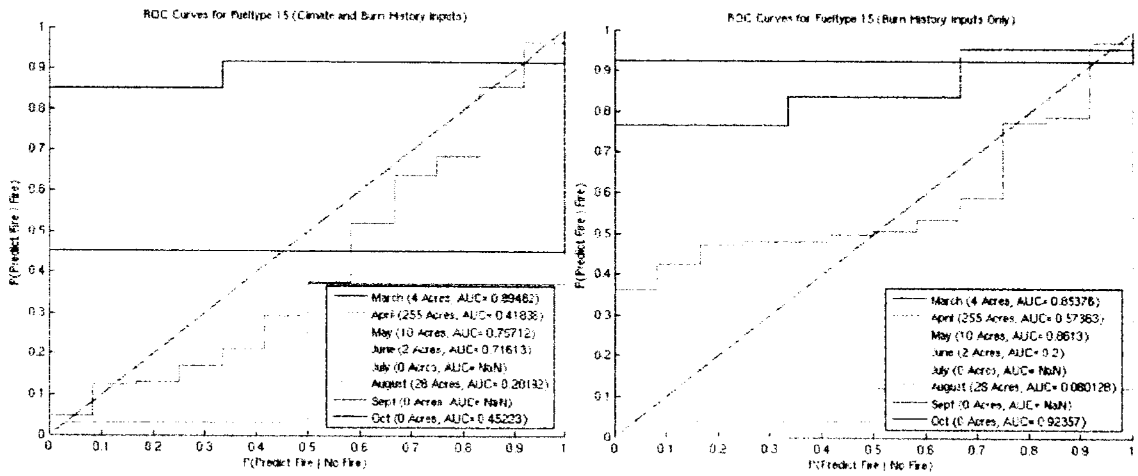
Fuel Code H: Short Needle Conifers, Normal Dead Load (logit method)



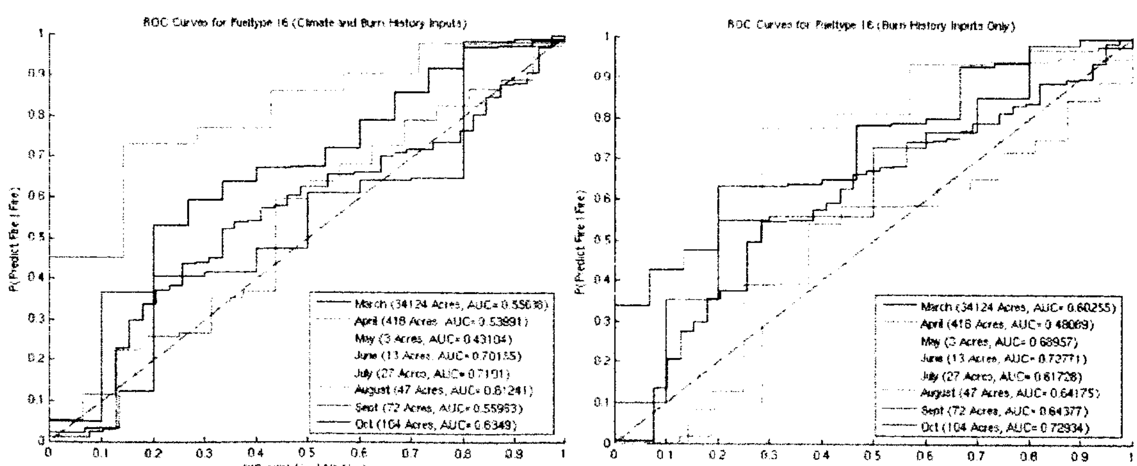
Fuel Code L: Western Perennial Grass (logit method)



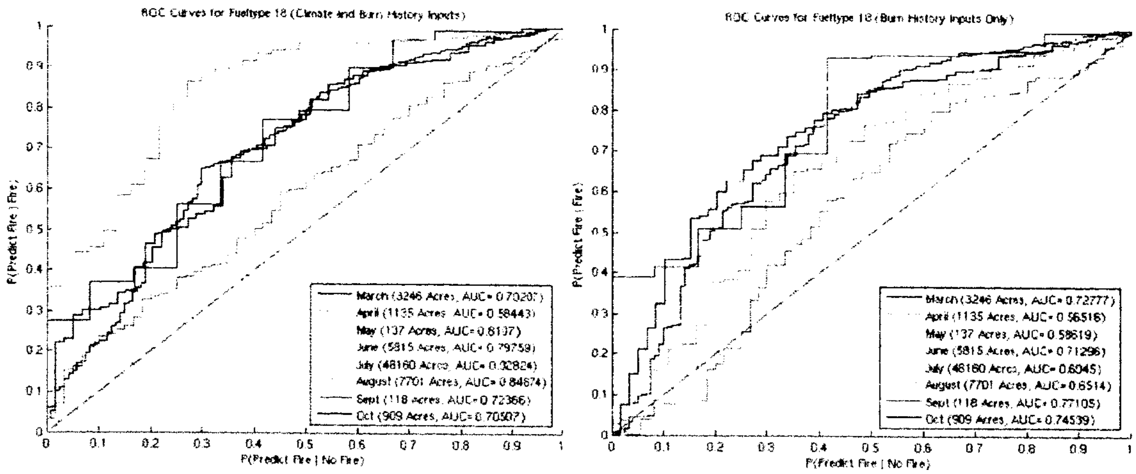
Fuel Code M: Agricultural Land (logit method)



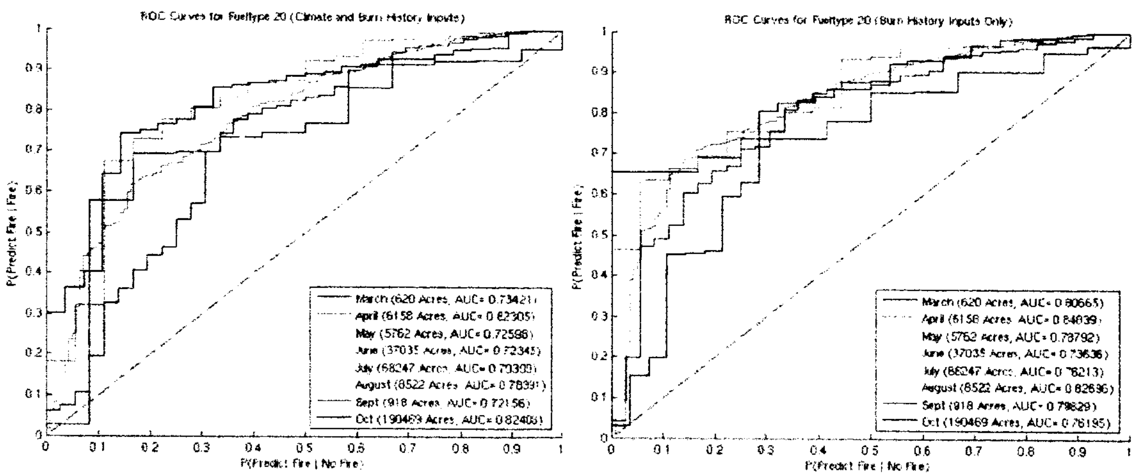
Fuel Code O: High Pocosin (logit method)



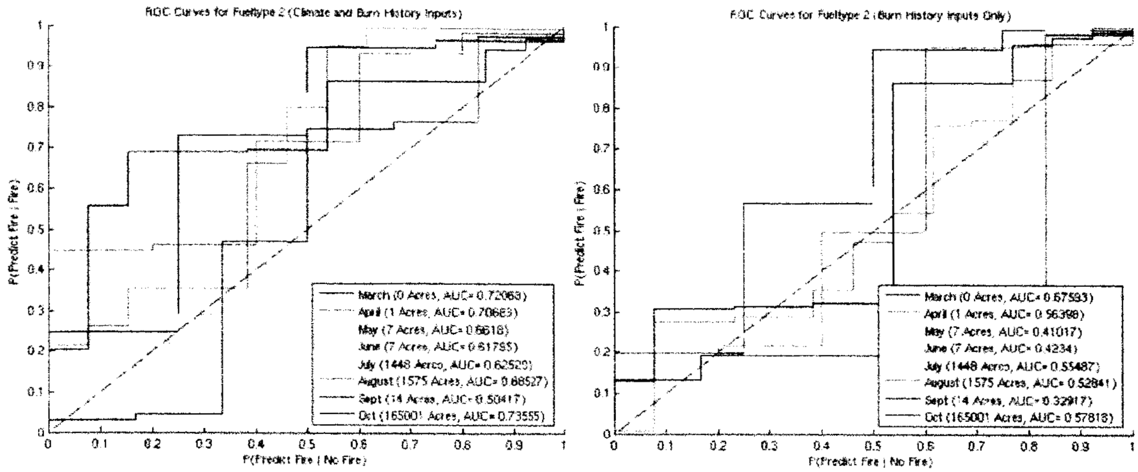
Fuel Code P: Southern Pine Plantation (logit method)



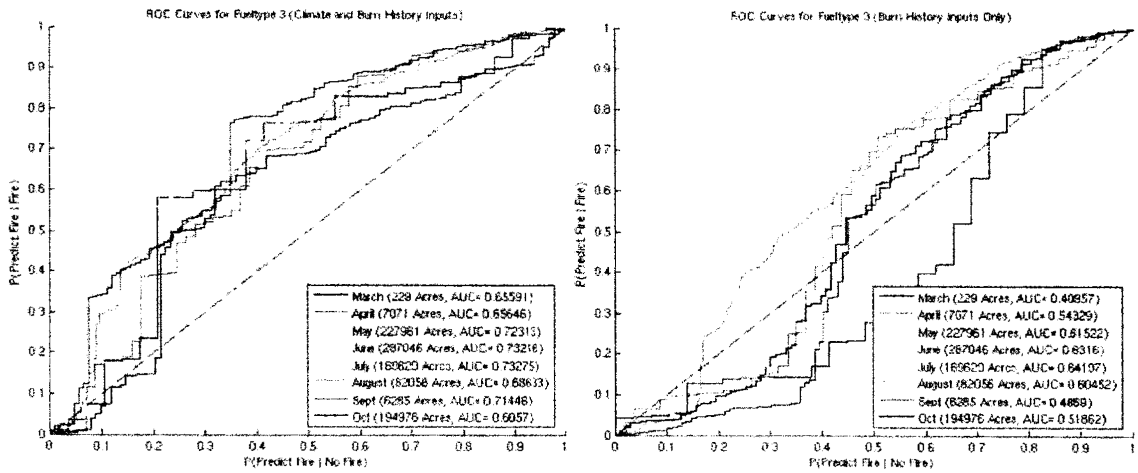
Fuel Code R: Hardwoods (logit method)



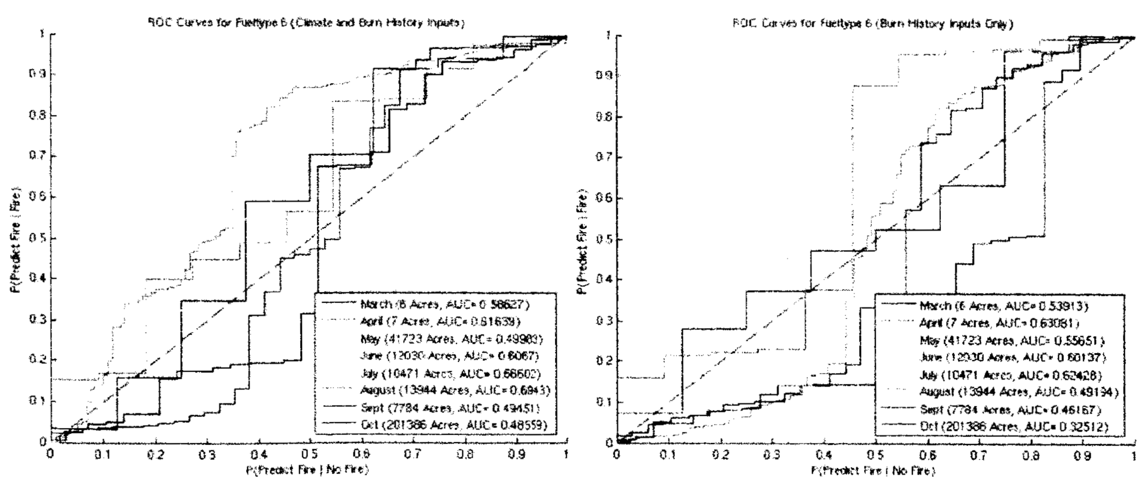
Fuel Code T: Sagebrush-Grass Mixture (logit method)



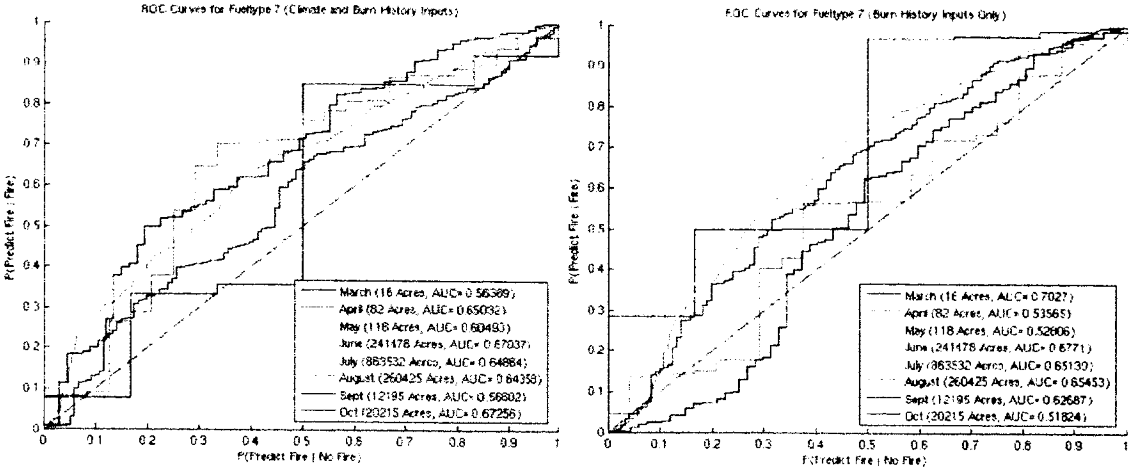
Fuel Code B: California Mixed Chaparral (CART method)



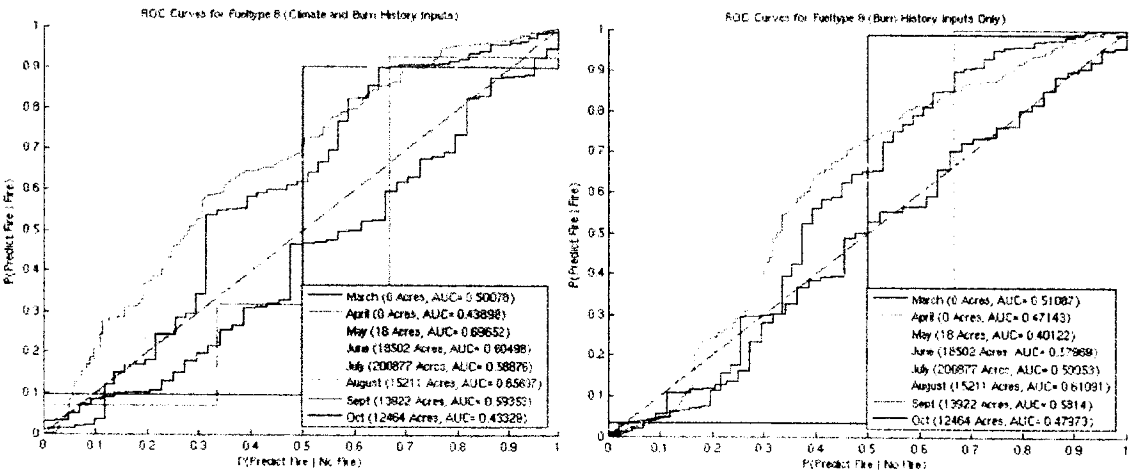
Fuel Code C: Pine Grass Savanna (CART method)



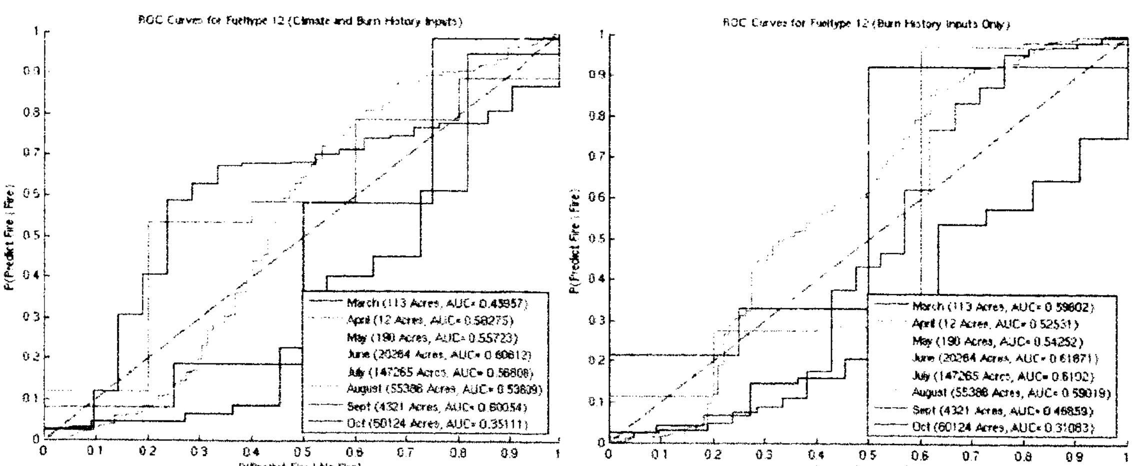
Fuel Code F: Intermediate Brush (CART method)



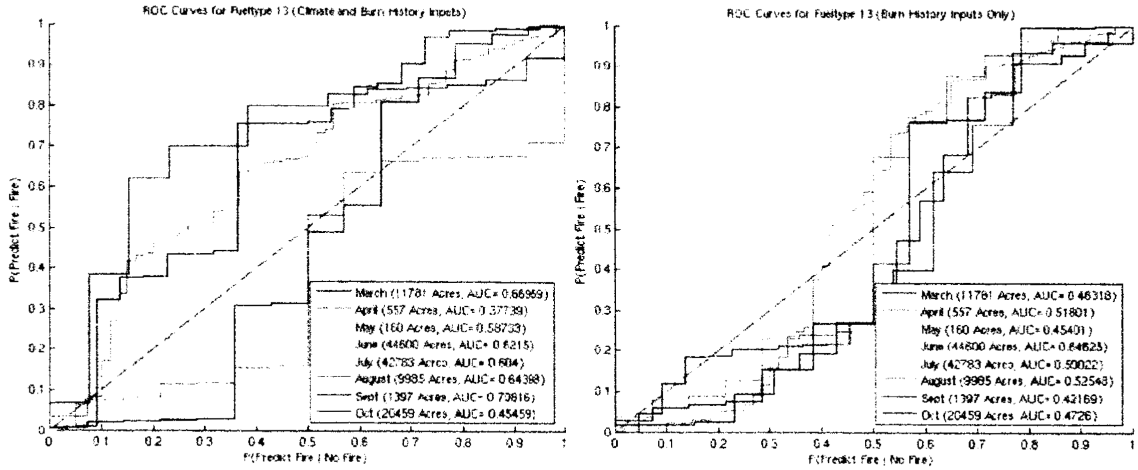
Fuel Code G: Short Needle Conifers, Heavy Dead Load (CART method)



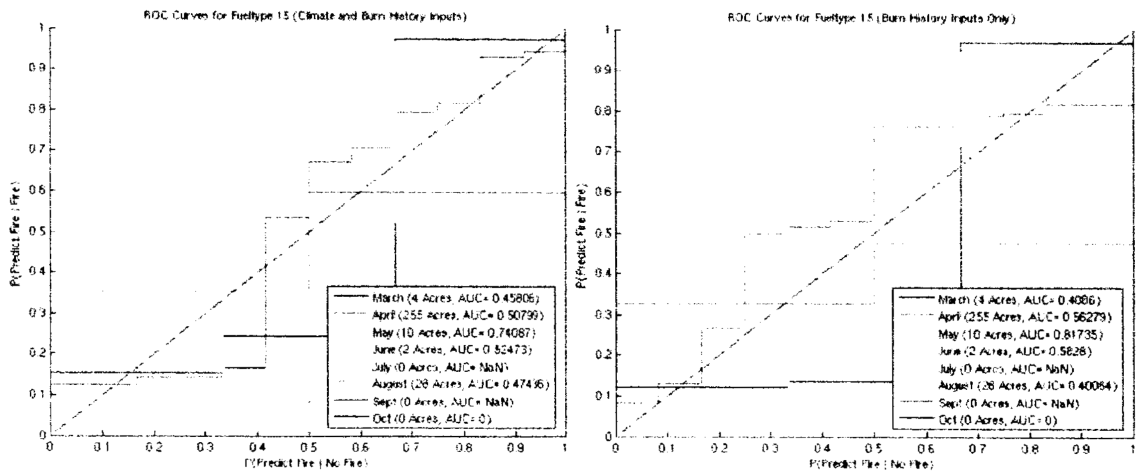
Fuel Code H: Short Needle Conifers, Normal Dead Load (CART method)



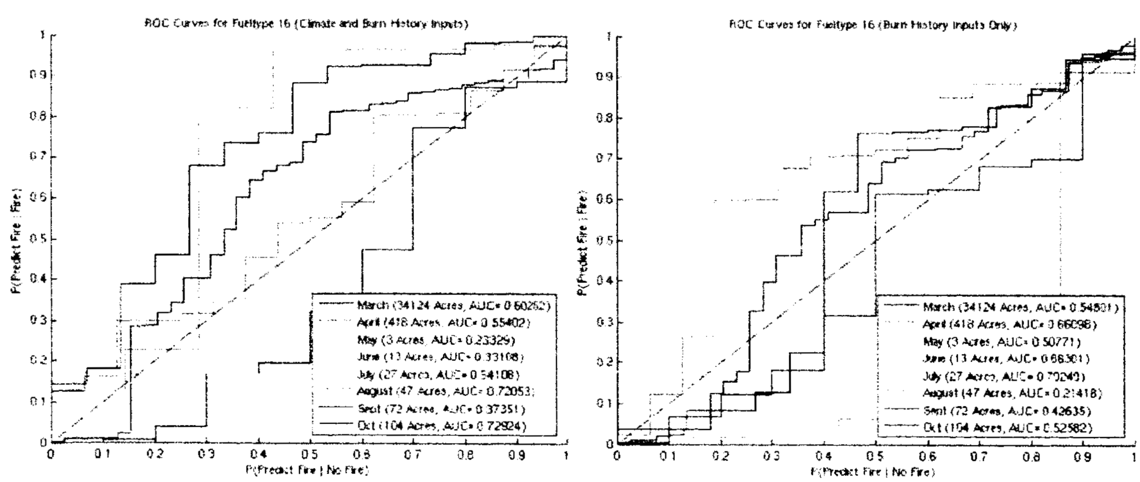
Fuel Code L: Western Perennial Grass (CART method)



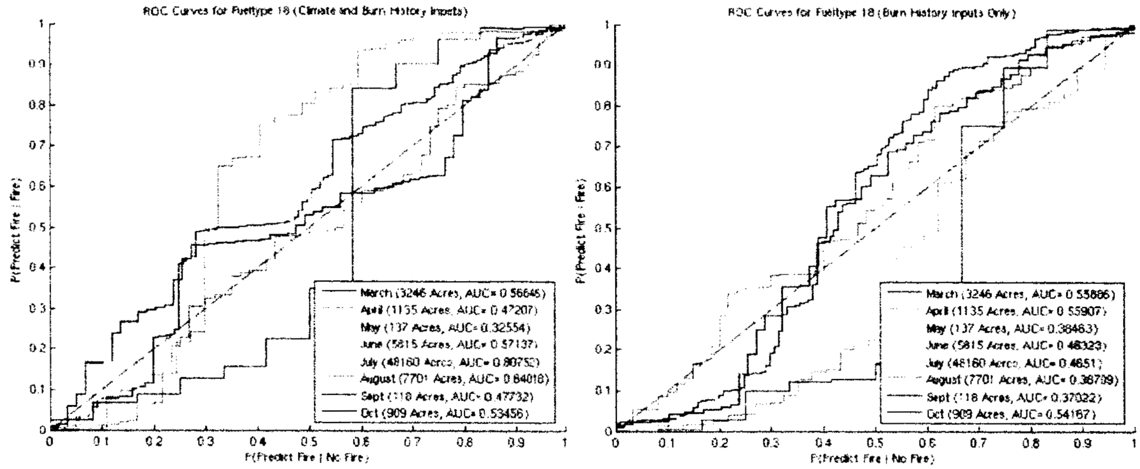
Fuel Code M: Agricultural Land (CART method)



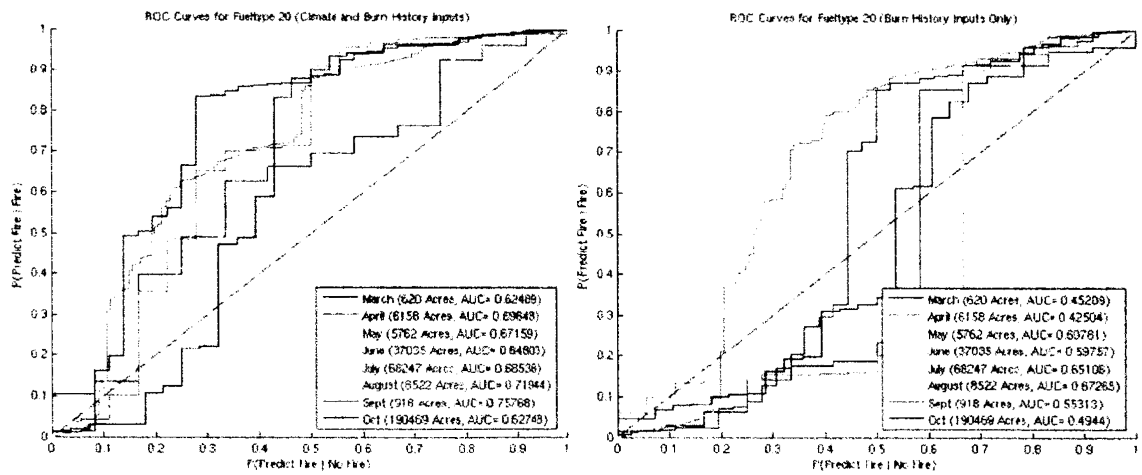
Fuel Code O: High Pocosin (CART method)



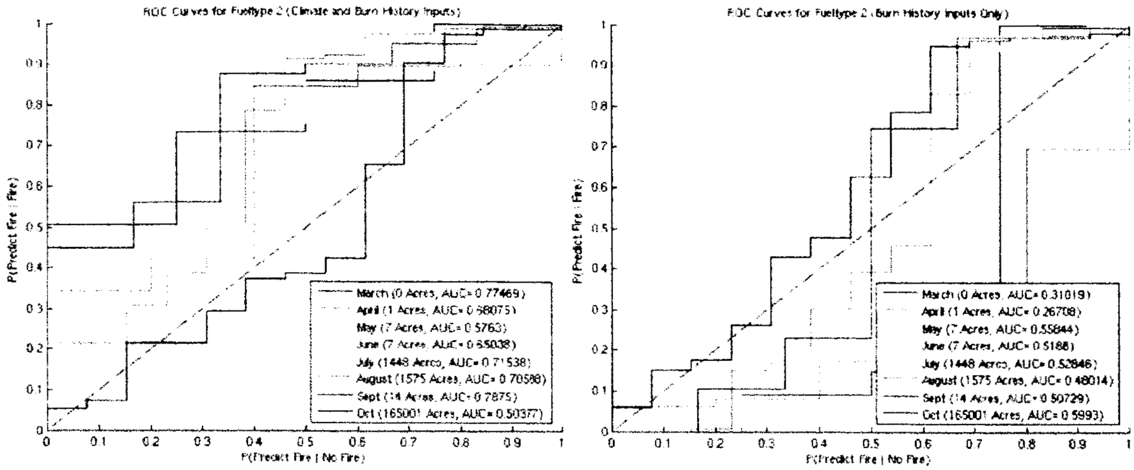
Fuel Code P: Southern Pine Plantation (CART method)



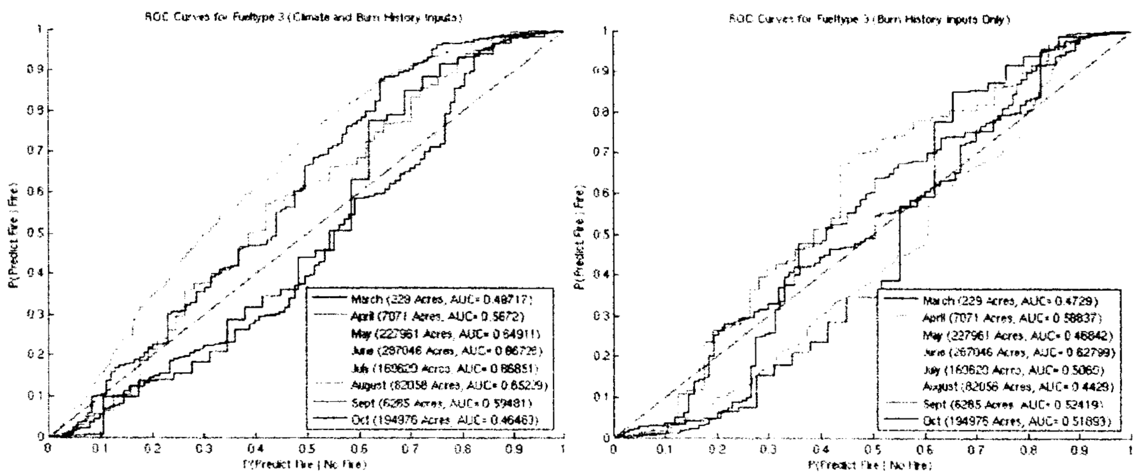
Fuel Code R: Hardwoods (CART method)



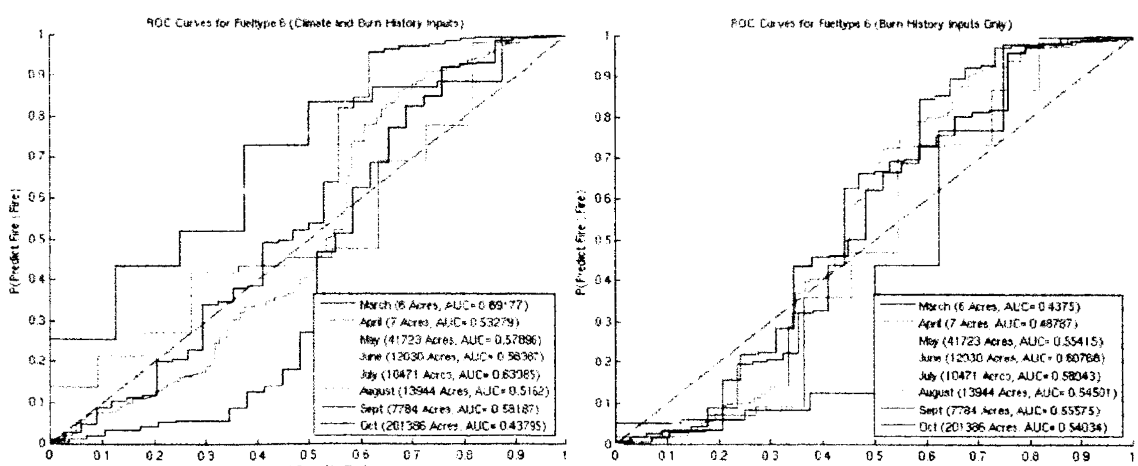
Fuel Code T: Sagebrush-Grass Mixture (CART method)



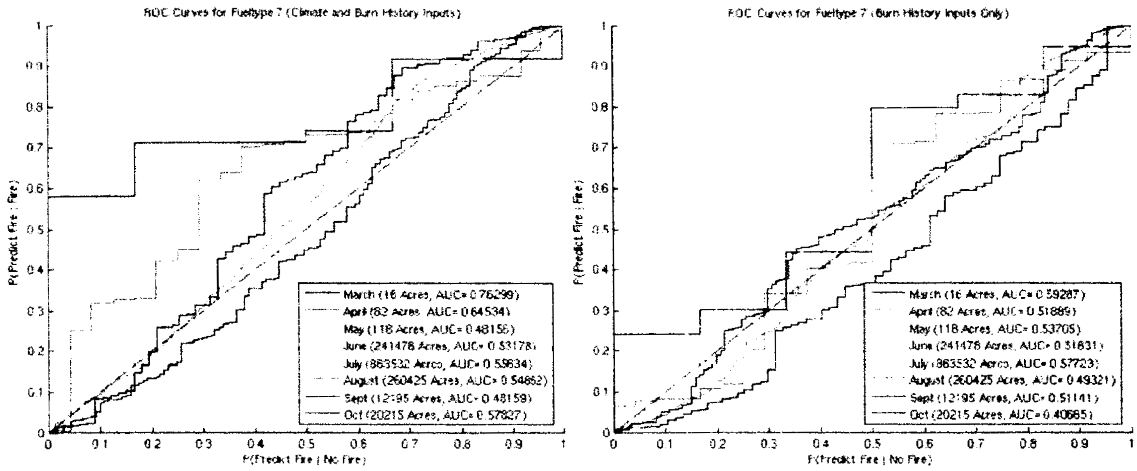
Fuel Code B: California Mixed Chaparral (SVM method)



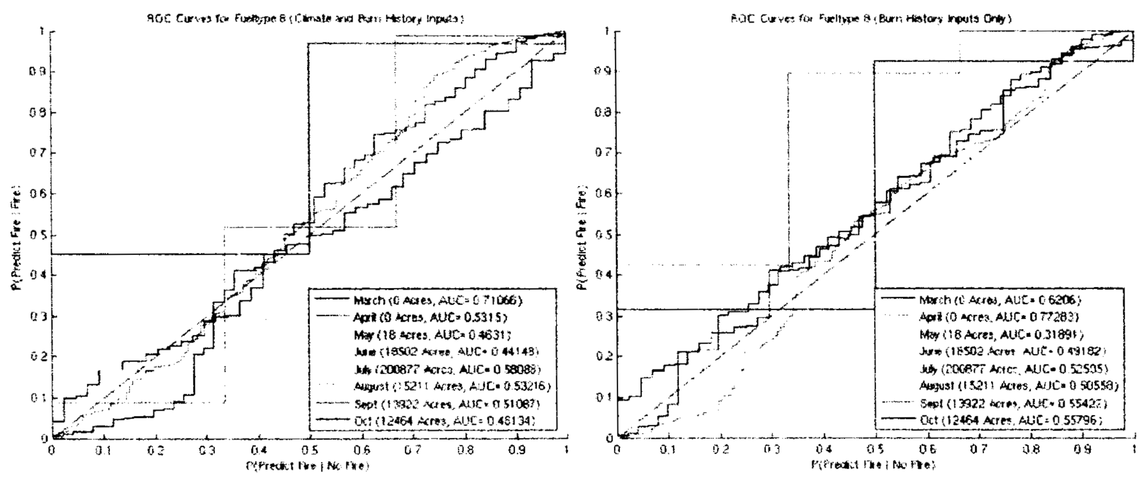
Fuel Code C: Pine Grass Savanna (SVM method)



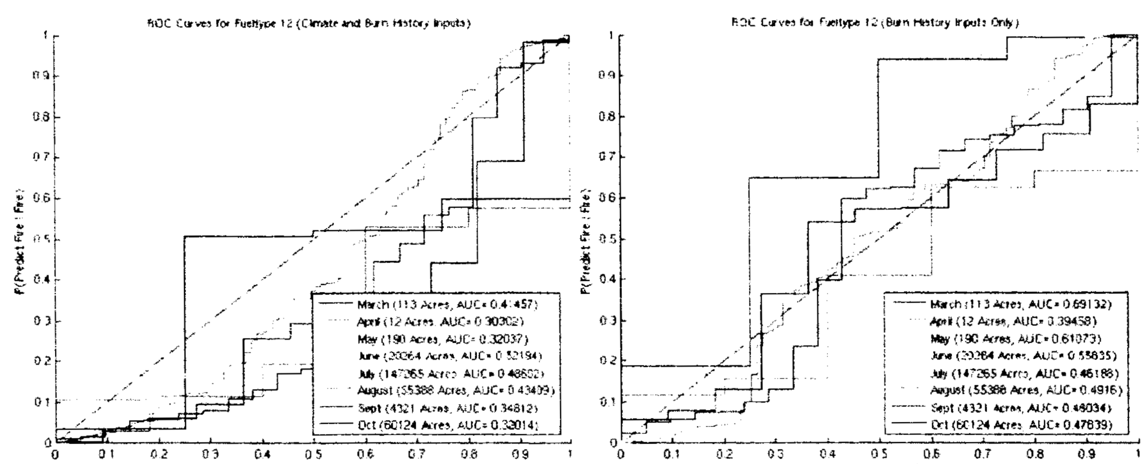
Fuel Code F: Intermediate Brush (SVM method)



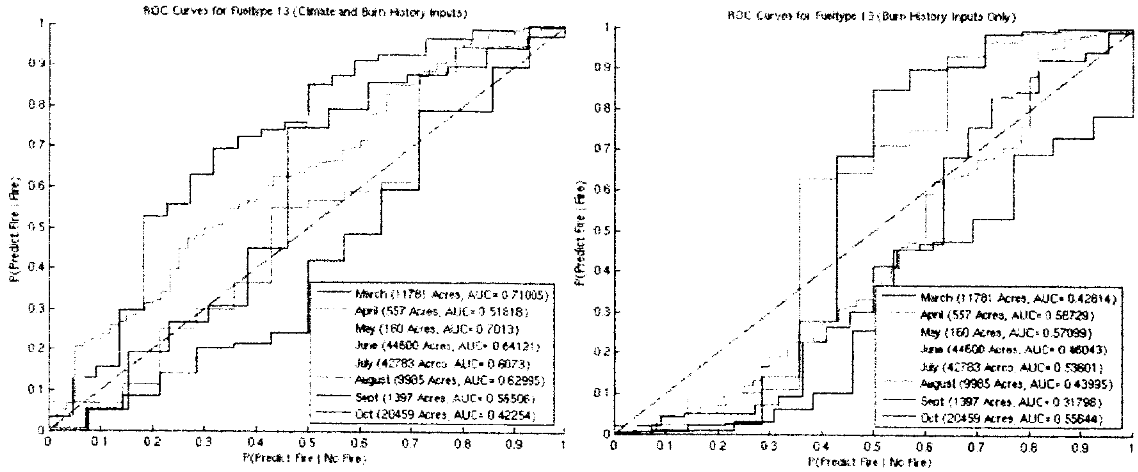
Fuel Code G: Short Needle Conifers, Heavy Dead Load (SVM method)



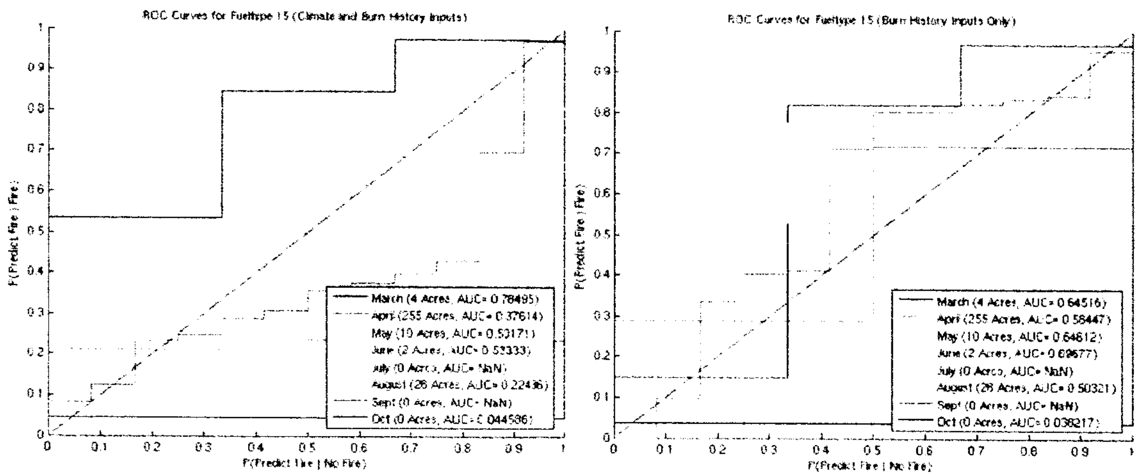
Fuel Code H: Short Needle Conifers, Normal Dead Load (SVM method)



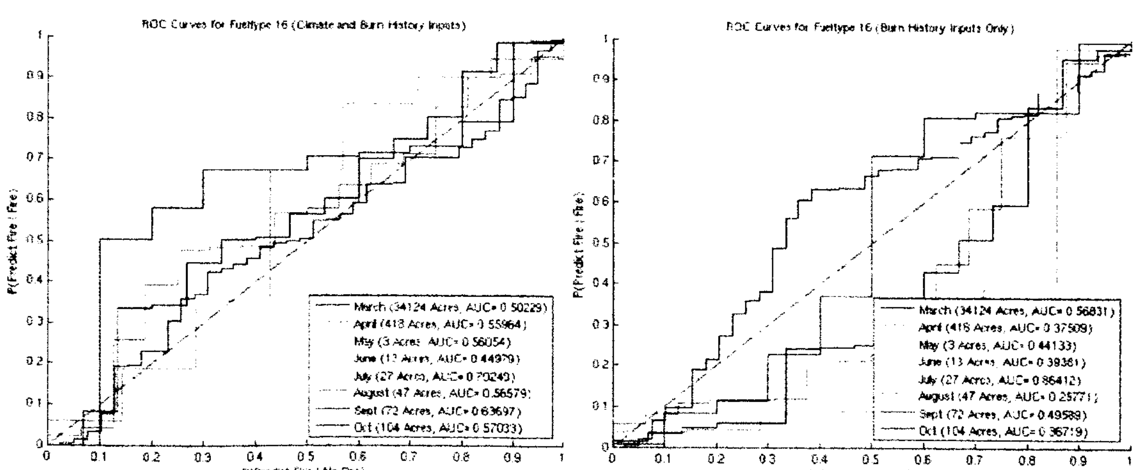
Fuel Code L: Western Perennial Grass (SVM method)



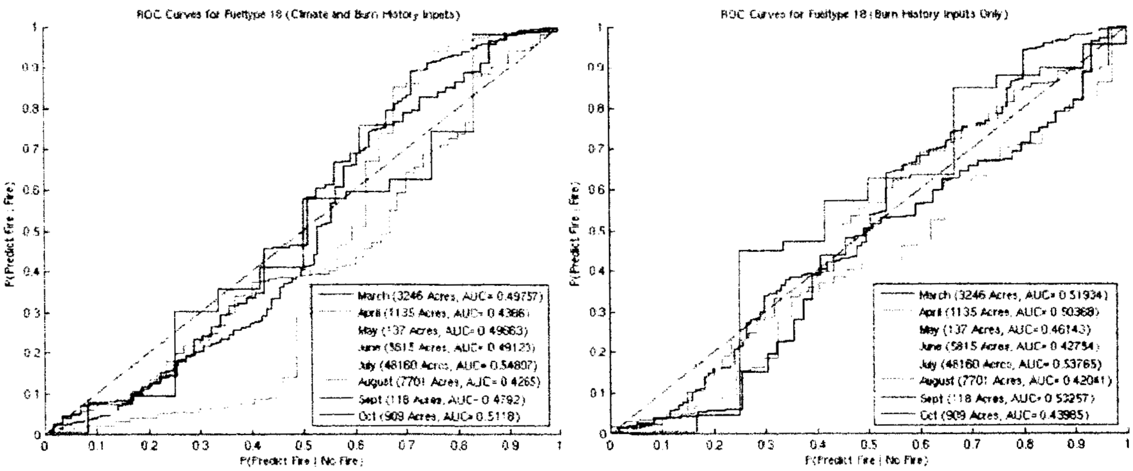
Fuel Code M: Agricultural Land (SVM method)



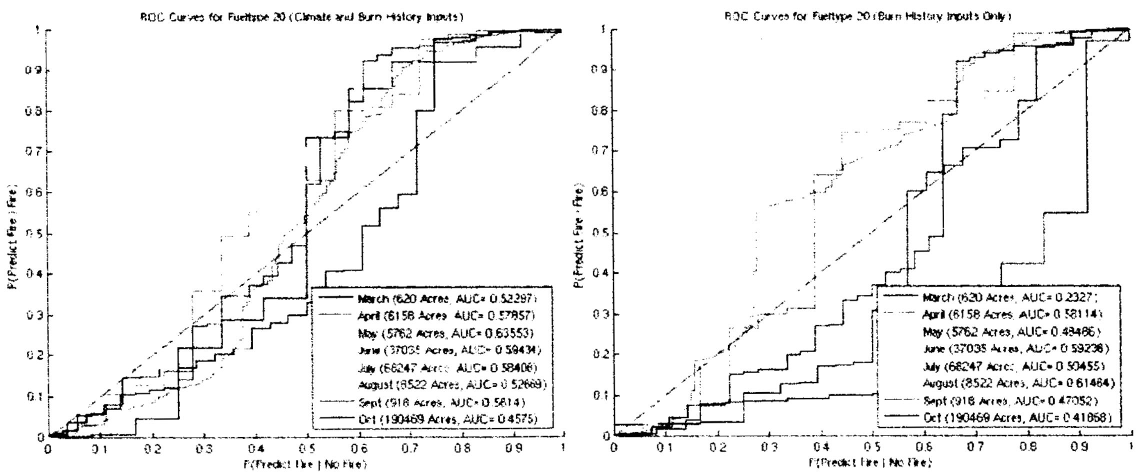
Fuel Code O: High Pocosin (SVM method)



Fuel Code P: Southern Pine Plantation (SVM method)



Fuel Code R: Hardwoods (SVM method)



Fuel Code T: Sagebrush-Grass Mixture (SVM method)

Appendix 2

Fuel Code	March	April	May	June	July	Aug	Sept	Oct
B			0.862 0.835	0.68 0.7664	0.716 0.7964	0.732 0.736		0.7587 0.7179
C	0.723 0.8235	0.7713 0.7493	0.825 0.7832	0.7565 0.7545	0.7513 0.7221	0.7279 0.7267	0.7944 0.8031	0.6544 0.7119
F		0.7368 0.8115	0.6776 0.728	0.6757 0.7046	0.7773 0.7858	0.828 0.811	0.755 0.7599	0.837 0.7111
G		0.7257 0.6258	0.6709 0.6775	0.7173 0.7269	0.6929 0.687	0.7022 0.6982	0.6313 0.6907	0.6816 0.7038
H			0.795 0.7783	0.6677 0.6716	0.6772 0.6687	0.716 0.7039	0.6307 0.7073	0.5155 0.5295
L			0.8438 0.8208	0.6894 0.7439	0.6128 0.6787	0.6948 0.6783	0.706 0.7537	0.6261 0.5832
M	0.6295 0.6509	0.7378 0.7806	0.6981 0.7368	0.722 0.7486	0.7662 0.7333	0.742 0.723	0.6925 0.7	0.7494 0.7618
O		0.4184 0.5736	0.7671 0.8613					
P	0.5564 0.6025	0.5399 0.4809					0.5596 0.6438	0.6349 0.7293
R	0.7021 0.7278	0.5844 0.5652	0.8107 0.5862	0.7976 0.713	0.9282 0.6945	0.8487 0.6514	0.7237 0.7711	0.7051 0.7454
T	0.7342 0.8067	0.823 0.8484	0.726 0.7879	0.7235 0.7364	0.794 0.7621	0.7839 0.827	0.7216 0.7983	0.8241 0.762

Table 1: Logistic Regression, All Inputs vs. Logistic Regression, Burn History Inputs

Fuel Code	March	April	May	June	July	Aug	Sept	Oct
B			0.6618 0.4102	0.618 0.4234	0.6253 0.5549	0.6853 0.5284		0.7355 0.5782
C	[illegible] [illegible]	0.6565 0.5433	0.7231 0.6152	0.7322 0.6316	0.7328 0.642	0.6863 0.6045	0.7145 0.4869	0.6057 0.5186
F		0.6164 0.6308	0.4998 0.5565	0.6067 0.6014	0.666 0.6243	0.6943 0.4919	0.4945 0.4617	0.4856 0.3251
G		0.6503 0.5356	0.6049 0.5281	0.6704 0.6771	0.6486 0.6514	0.6436 0.6545	0.566 0.6269	0.6726 0.5182
H			0.6965 0.4012	0.605 0.5797	0.5888 0.5995	0.657 0.6109	0.5935 0.5814	0.4333 0.4797
L			0.5572 0.5425	0.6081 0.6187	0.5681 0.6192	0.5381 0.5902	0.6005 0.4686	0.3511 0.3108
M	0.6697 0.4632	0.3774 0.518	0.5873 0.454	0.6215 0.6463	0.604 0.5902	0.644 0.5255	0.7082 0.4217	0.4546 0.4726
O		0.508 0.5628	0.7409 0.8174					
P	0.6026 0.548	0.554 0.661					0.3735 0.4263	0.7292 0.5258
R	0.5665 0.5589	0.4721 0.5591	0.3255 0.3846	0.5714 0.4832	0.8075 0.4651	0.6402 0.3871	0.4773 0.3702	0.5346 0.5417
T	0.6249 0.4521	0.6965 0.425	0.6716 0.6078	0.648 0.5976	0.6854 0.6511	0.7194 0.6726	0.7577 0.5531	0.6275 0.4944

Table 2: CART, All Inputs vs. CART, Burn History Inputs

Fuel Code	March	April	May	June	July	Aug	Sept	Oct
B			0.5763 0.5584	0.6504 0.5188	0.7154 0.5285	0.7059 0.4801		0.5038 0.5993
C	0.4972 0.4729	0.5672 0.5884	0.6491 0.4684	0.6673 0.628	0.6685 0.5069	0.6523 0.4429	0.5948 0.5242	0.4646 0.5189
F		0.5328 0.4879	0.579 0.5541	0.5839 0.6077	0.6398 0.5894	0.5162 0.545	0.5819 0.5558	0.4379 0.5403
G		0.6453 0.5189	0.4816 0.537	0.5318 0.5163	0.5563 0.5772	0.5486 0.4932	0.4816 0.5114	0.5783 0.4064
H			0.4631 0.3189	0.4415 0.4918	0.5809 0.5259	0.5322 0.5056	0.5109 0.5542	0.4813 0.558
L			0.3204 0.6107	0.5219 0.5584	0.4869 0.4619	0.4341 0.4916	0.3481 0.4803	0.3201 0.4784
M	0.7101 0.4281	0.5182 0.5673	0.7013 0.571	0.6412 0.4604	0.6073 0.536	0.6299 0.44	0.5651 0.318	0.4225 0.5564
O		0.3761 0.5845	0.5017 0.6461					0.0446 0.0382
P	0.5023 0.5683	0.5596 0.3751					0.637 0.4959	0.5703 0.3672
R	0.4976 0.5193	0.4366 0.5037	0.4966 0.4614	0.4912 0.4275	0.549 0.5376	0.4265 0.4204	0.4792 0.5326	0.5118 0.4399
T	0.523 0.2327	0.5786 0.5811	0.6355 0.4849	0.5943 0.5924	0.5841 0.5946	[redacted]	0.5614 0.4705	0.4575 0.4187

Table 3: SVM, All Inputs vs. SVM, Burn History Inputs

Fuel Code	March	April	May	June	July	Aug	Sept	Oct
B			0.862 0.6618	0.68 0.618	0.716 0.6253	0.732 0.6853		0.7587 0.7355
C	0.723 0.6559	0.7713 0.6565	0.825 0.7231	0.7565 0.7322	0.7513 0.7328	0.7279 0.6863	0.7944 0.7145	0.6544 0.6057
F		0.7368 0.6164	0.6776 0.4998	0.6757 0.6067	0.7773 0.666	0.828 0.6941	0.755 0.4945	0.837 0.4856
G		0.7257 0.6503	0.6709 0.6049	0.7173 0.6704	0.6929 0.6486	0.7022 0.6436	0.6313 0.566	0.6816 0.6726
H			0.795 0.6965	0.6677 0.605	0.6772 0.5888	0.716 0.657	0.6307 0.5935	0.5155 0.4333
L			0.8438 0.5572	0.6894 0.6081	0.6128 0.5681	0.6948 0.5381	0.706 0.6005	0.6261 0.3511
M	0.6295 0.6697	0.7378 0.3774	0.6981 0.5873	0.722 0.6215	0.7662 0.604	0.742 0.644	0.6925 0.7082	0.7494 0.4546
O		0.4184 0.508	0.7671 0.7409					
P	0.5564 0.6026	0.5399 0.554					0.5596 0.3735	0.6349 0.7292
R	0.7021 0.5665	0.5844 0.4721	0.8107 0.3255	0.7976 0.5714	0.9282 0.8075	0.8487 0.6402	0.7237 0.4773	0.7051 0.5346
T	0.7342 0.6249	0.823 0.6965	0.726 0.6716	0.7235 0.648	0.794 0.6854	0.7839 0.7194	0.7216 0.7577	0.8241 0.6275

Table 4: Logistic Regression, All Inputs vs. CART, All Inputs

Fuel Code	March	April	May	June	July	Aug	Sept	Oct
B			0.862 0.5763	0.68 0.6504	0.716 0.7154	0.732 0.7059		0.7587 0.5038
C	0.729 0.4972	0.7713 0.5672	0.825 0.6491	0.7565 0.6673	0.7513 0.6685	0.7279 0.6523	0.7944 0.5948	0.6544 0.4646
F		0.7368 0.5328	0.6776 0.579	0.6757 0.5839	0.7773 0.6398	0.828 0.5162	0.755 0.5819	0.837 0.4379
G		0.7257 0.6453	0.6709 0.4816	0.7173 0.5318	0.6929 0.5563	0.7022 0.5486	0.6313 0.4816	0.6816 0.5783
H			0.785 0.4631	0.6677 0.4415	0.6772 0.5809	0.716 0.5322	0.6367 0.5109	0.5155 0.4813
L			0.8438 0.3204	0.6894 0.5219	0.6128 0.4869	0.6948 0.4341	0.706 0.3481	0.6261 0.3201
M	0.6295 0.7101	0.7378 0.5182	0.6981 0.7013	0.722 0.6412	0.7662 0.6073	0.742 0.6299	0.6925 0.5651	0.7494 0.4225
O		0.4184 0.3761	0.7671 0.5017					
P	0.5564 0.5023	0.5399 0.5596					0.5596 0.637	0.6349 0.5703
R	0.7021 0.4976	0.5844 0.4366	0.8107 0.4966	0.7976 0.4912	0.9282 0.549	0.8487 0.4265	0.7237 0.4792	0.7051 0.5118
T	0.7342 0.523	0.823 0.5786	0.726 0.6355	0.7235 0.5943	0.794 0.5841	0.7839 0.5267	0.7216 0.5614	0.8241 0.4575

Table 5: Logistic Regression, All Inputs vs. SVM, All Inputs

Fuel Code	March	April	May	June	July	Aug	Sept	Oct
B			0.6618 0.5763	0.618 0.6504	0.6253 0.7154	0.6853 0.7059		0.7355 0.5038
C	0.6559 0.4972	0.6565 0.5672	0.7231 0.6491	0.7322 0.6673	0.7328 0.6685	0.6863 0.6523	0.7145 0.5948	0.6057 0.4646
F		0.6164 0.5328	0.4998 0.579	0.6067 0.5839	0.666 0.6398	0.6943 0.5162	0.4945 0.5819	0.4856 0.4379
G		0.6503 0.6453	0.6049 0.4816	0.6704 0.5318	0.6486 0.5563	0.6436 0.5486	0.566 0.4816	0.6726 0.5783
H			0.6965 0.4631	0.605 0.4415	0.5888 0.5809	0.657 0.5322	0.5935 0.5109	0.4333 0.4813
L			0.5572 0.3204	0.6081 0.5219	0.5681 0.4869	0.5381 0.4341	0.6005 0.3481	0.3511 0.3201
M	0.6697 0.7101	0.3774 0.5182	0.5873 0.7013	0.6215 0.6412	0.604 0.6073	0.644 0.6299	0.7082 0.5651	0.4546 0.4225
O		0.508 0.3761	0.7409 0.5017					
P	0.6026 0.5023	0.554 0.5596					[redacted]	0.7292 0.5703
R	0.5665 0.4976	0.4721 0.4366	[redacted]	0.5714 0.4912	0.8075 0.549	0.6402 0.4265	0.4773 0.4792	0.5346 0.5118
T	0.6249 0.523	0.6965 0.5786	0.6716 0.6355	0.648 0.5943	0.6854 0.5841	0.7194 0.5267	0.7577 0.5614	0.6275 0.4575

Table 6: CART, All Inputs vs. SVM, All Inputs