Biological Evaluation
R2-06-04

Biological Evaluation of a Model for Predicting Presence of White Pine Blister Rust in Colorado Based on Climatic Variables and Susceptible White Pine Species Distribution

Biological Evaluation
R2-06-04

Biological Evaluation of a Model for Predicting Presence of White Pine Blister Rust in Colorado Based on Climatic Variables and Susceptible White Pine Species Distribution

January 2006

Prepared by:

____
Brian Howell
Biological Technician, Lakewood Service Center (LSC), Forest Health Management

____
Kelly Burns
Pathologist, LSC, Forest Health Management

and

____
Holly Kearns
Pathologist, Coeur d'Alene Field Office, Forest Health Protection

Approved by:

____
Jeffrey J. Witcosky
Entomologist and Leader, LSC, Forest Health Management

and

____
Frank J. Cross
Group Leader, Forest Health Management

USDA Forest Service
Rocky Mountain Region
Renewable Resources
PO Box 25127
Lakewood, CO 80225-0127

Introduction

Kearns (2005) developed a classification and regression tree (CART) model to predict the presence or absence of white pine blister rust (WPBR) in stands containing susceptible white pine species in Colorado. This CART analysis employed the multinomial statistical model in which splits are based on minimizing the deviance, defined by the multinomial log-likelihood. The model was created using S-PLUS software. The model initially utilized both weather data based on 30-year monthly averages of weather variables taken from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) dataset (Daly et al. 2002) and plot data taken from 329 plots located throughout Colorado and Wyoming. When these variables were entered as potential predictors of the presence or absence of WPBR, certain weather variables were selected, but none of the plot variables were found to be significant predictors. The four weather variables found to be significant that were selected for splits in the classification tree were: May relative humidity, May precipitation, May minimum temperature and August minimum temperature (Figure 1).

Kearns (2005) applied this predictive model to four vegetation coverage data sources that map the distribution of white pines on the landscape. This was done using ArcGIS software by first locating the centroid of each white pine polygon from each of the data sources. This centroid was assumed to be representative of conditions in each polygon and the model was run by evaluating the weather conditions pertinent to the model at the centroid point and following the model's decision tree to an outcome of either presence or absence of WPBR. Once the model found the centroid to indicate presence or absence, the entire polygon was labeled reflecting that decision.

This was a logical methodology that allowed for the comparison of differences between data sources, but produced some problematic results. When all four of the data sources were examined simultaneously on a map, polygons from different data sources tended to overlap. Occasionally, where this occurred, contradictory results were displayed; i.e. the same parcel of land was labeled as both having a presence and an absence of WPBR (Figures 2 and 5). It appears that one of Kearn's (2005) intentions was to evaluate the differences between data sources and for this purpose the contradictory results were useful. For our purpose, which is to provide land managers with a tool to evaluate risk of WPBR establishment, the contradictory results were difficult to interpret.

To avoid confusion associated with contradictory results, and to optimize the precision of the model output, the model was re-run looking first only at the weather variables, ignoring the data sources that predict the presence of white pine on the landscape. The rationale for this method was that the limit to the precision of the model output is the resolution of the data that it is based on, in this case the 2-4 km resolution of the PRISM weather data. The model was run for the entire the state of Colorado, the majority of which is not suitable for white pine establishment. However, this broad coverage is irrelevant as the model is intended to be used only where white pine occurs, so results from areas without white pine should be ignored. Once the model output represented the highest attainable precision (Figures 3 and 6), land managers could overlay whichever white pine data source they favor for their local area (Figure 4) and be confident that the resulting predictions were based on the best available data. The selected vegetation coverage dataset may be one of the ones used by Kearns (2005) or an updated and more accurate coverage not used in the initial analysis.

Methods

Kearns (2005) described in detail the methodology for creation of the CART model predicting presence or absence of WPBR (Figure 1). Here we describe only the methods for re-running this model based on the weather data.

The datasets containing the 30-year monthly averages for the weather variables (calculated from the PRISM dataset) that were originally used by Kearns (2005) were provided by Colorado State University and the USDA Forest Service Forest Health Technology Enterprise Team. These datasets consisted of either 2 or 4 km resolution raster files. Each 2 or 4 km square cell had a value associated with it which represented the 30-year monthly average for the variable that was represented by the dataset. To reduce the complexity of these datasets to the minimum level required to run this model, the values of all the cells were reclassified as shown in Table 1 based on the values of the splits in the CART model (Figure 1). Each of the four splits was assigned its own place value such that once the reassigned values were summed for individual polygons, each place value could be evaluated individually to follow the CART decision tree model. For example, the outcome 2212 can be followed down the decision tree where 2xxx indicates May relative humidity $\geq$ 54.5%, x2xx indicates May precipitation $\geq$ 87.875 mm, and xx1x indicates August minimum temperature < 8.79 C at which point the decision tree indicates absence, regardless of whether the outcome is 2212 or 2211 (Table 2 and Figure 1).

Three of the four variables, May precipitation, May minimum temperature, and August minimum temperature, had a resolution of 4 km and the cell boundaries for each of these datasets coincided perfectly, i.e. each cell boundary was in exactly the same location for each of these datasets. The fourth variable, May relative humidity, had a resolution of 2 km and the boundaries from this dataset did not coincide with the other three. To maintain the resolution of each of the original raster datasets, each was converted to a vector format using ArcToolbox software and all of the datasets were spatially joined into a single shapefile that contained attributes from all of the raster datasets. The function “identity” in ArcToolbox was used to perform the spatial joins. The resulting shapefile contained cells of various sizes due to the mismatch between the 4 km grids and the 2 km grid, however each cell contained spatially accurate data for each of the four reclassified weather variables based on the original raster data files. To run the model the four reclassified values (Table 1) were added for each cell and the sum was evaluated to indicate presence or absence (Table 2), by following the classification tree produced by the CART model (Figure 1). The final column of the attribute table (entitled “P_A”) of the model output shape file contains the output of the model where a “0” indicates absence and a “1” indicates presence of WPBR.

Table 1. Values of the splits for the reclassification of the original values from each of the raster-based PRISM weather datasets and their reclassified values utilized for re-running the Kearns (2005) WPBR CART model.

Variable	Original Value	Reclassified Value
May Relative Humidity	< 54.5 %	1000
	≥ 54.5 %	2000
May Precipitation	< 87.875 mm	100
	≥ 87.875 mm	200
August Minimum Temperature	< 8.79° C	10
	≥ 8.79° C	20
May Minimum Temperature	< -0.79° C	1
	≥ -.079° C	2

Table 2. Sixteen possible outcomes resulting from the addition of the reclassified weather data and their classification as indicating presence or absence based on the Kearns (2005) WPBR CART model.

Presence	Absence
1112	1111
1212	1211
1122	1121
1222	1221
2111	2211
2121	2212
2112	
2122	
2221	
2222	

Applications

This model is designed to predict where WPBR will and will not occur as inoculum becomes more widespread with continued spread of the pathogen. Currently, WPBR is found in northern Colorado near Red Feather Lakes and in southern Colorado in the Sangre de Cristo and Wet Mountains (Figure 8), and recent surveys have indicated that the disease continues to spread into previously uninfested stands.

With the output of this model, land managers can use the best available vegetation coverage to develop a predictive map. Many attempts have been made at classifying vegetation cover types and many of these attempts have been criticized for their lack of reliability in practical application. Generally, land managers are familiar with various coverages and their local reliability, for this reason, we have left the decision of which vegetation coverage to use in conjunction with the model up to the local manager. Once a decision has been made regarding the most reliable coverage, a risk map can be produced by combining data from the vegetation coverage and the model output (Figures 4 and 7). The basic steps required to produce this type of map are as follows:

1) Construct a shapefile that contains only polygons from the vegetation coverage with the presence of susceptible white pine species, in Colorado these include: southwestern white pine (*Pinus strobiformis*), limber pine (*P. flexilis*), and Rocky Mountain bristlecone pine (*P. aristata*).

2) Join this shapefile with the model output shapefile (found at http://www.fs.fed.us/r2/fhm/downloads/ then select "wpbr_co_mdloutpt") using the "intersect" function in ArcToolbox with the model output shapefile as the "input feature" and the white pine vegetation coverage as the "intersect feature". This will clip the model output to the extent of the white pine polygons, while retaining the resolution of the model output. Individual white pine polygons are likely to be cut into smaller polygons based on the boundaries between the model's predictions of presence and absence.

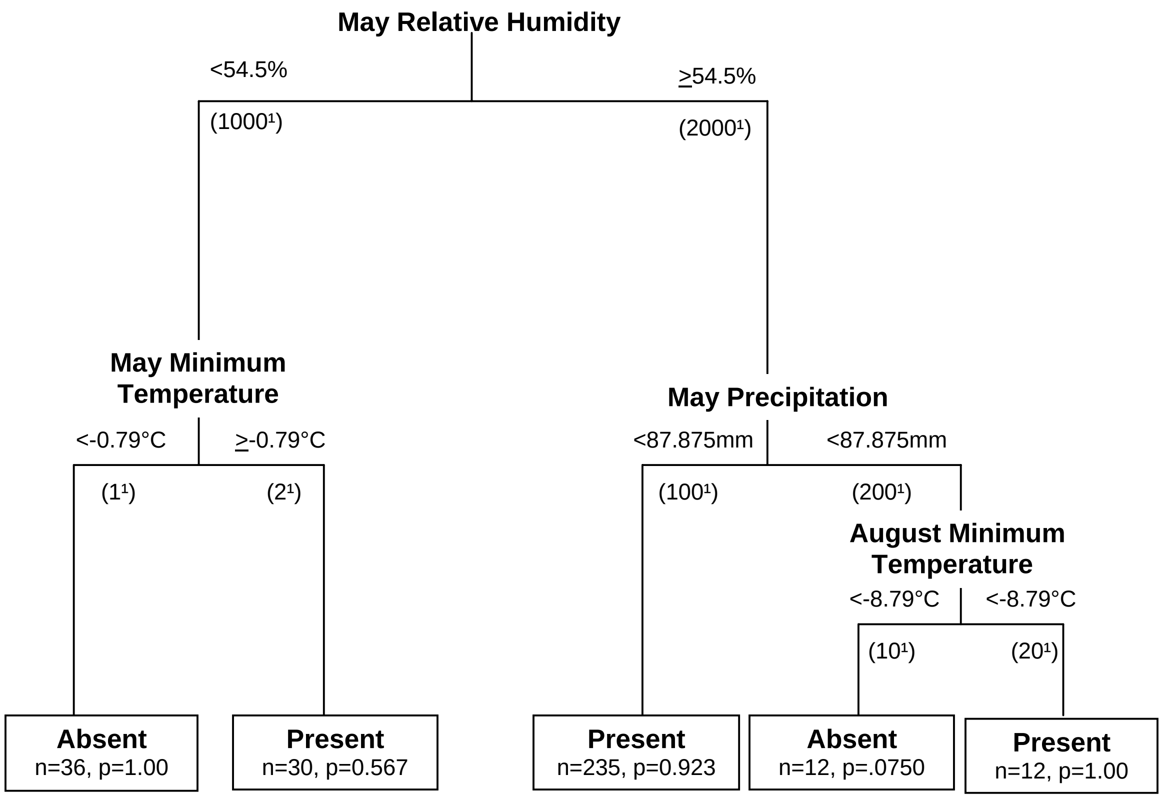
3) To display presence and absence on a map, double click on the output shapefile from #2 above. Select the "Symbology" tab. Select "Categories" in the "Show" box. Select "P_A" in the "Value Field" dropdown box. Select "Add All Values" and "OK". The two values displayed will be "0" and "1" where "0" indicates a prediction of absence and "1" indicates a prediction of presence of WPBR.

This model output can also be used to evaluate the WPBR hazard for known white pine stands not found in vegetation coverages. These are areas where white pine is observed on the ground, but no vegetation model predicts its presence. In this case, the best way to use the model is to complete the steps outlined in #3 above for the entire model output (figures 3 and 6; this will produce some interesting results- for example, sage and PJ cover types are generally predicted by the model to have a presence of WPBR- just ignore this). Locate the area known to have white pine on the map and examine the model output in that area to determine its prediction of presence or absence of WPBR.

Literature Cited

Daly, C., Gibson, W.P., Taylor, G.H., Johnson, G.L., and Pasteris, P. 2002. A knowledge-based approach to the statistical mapping of climate. Climate Research 22: 99-113.

Kearns, H.S.J. 2005. White pine blister rust in the central Rocky Mountains: Modeling current status and potential impacts (Ph.D. diss., Colorado State University, Fort Collins, Colorado). 243 p.



[1] Values as reclassified for the purpose of simplifying re-running the model.

Figure 1. Classification tree for presence/absence of white pine blister rust based on both plot and meteorological variables (from Kearns 2005).

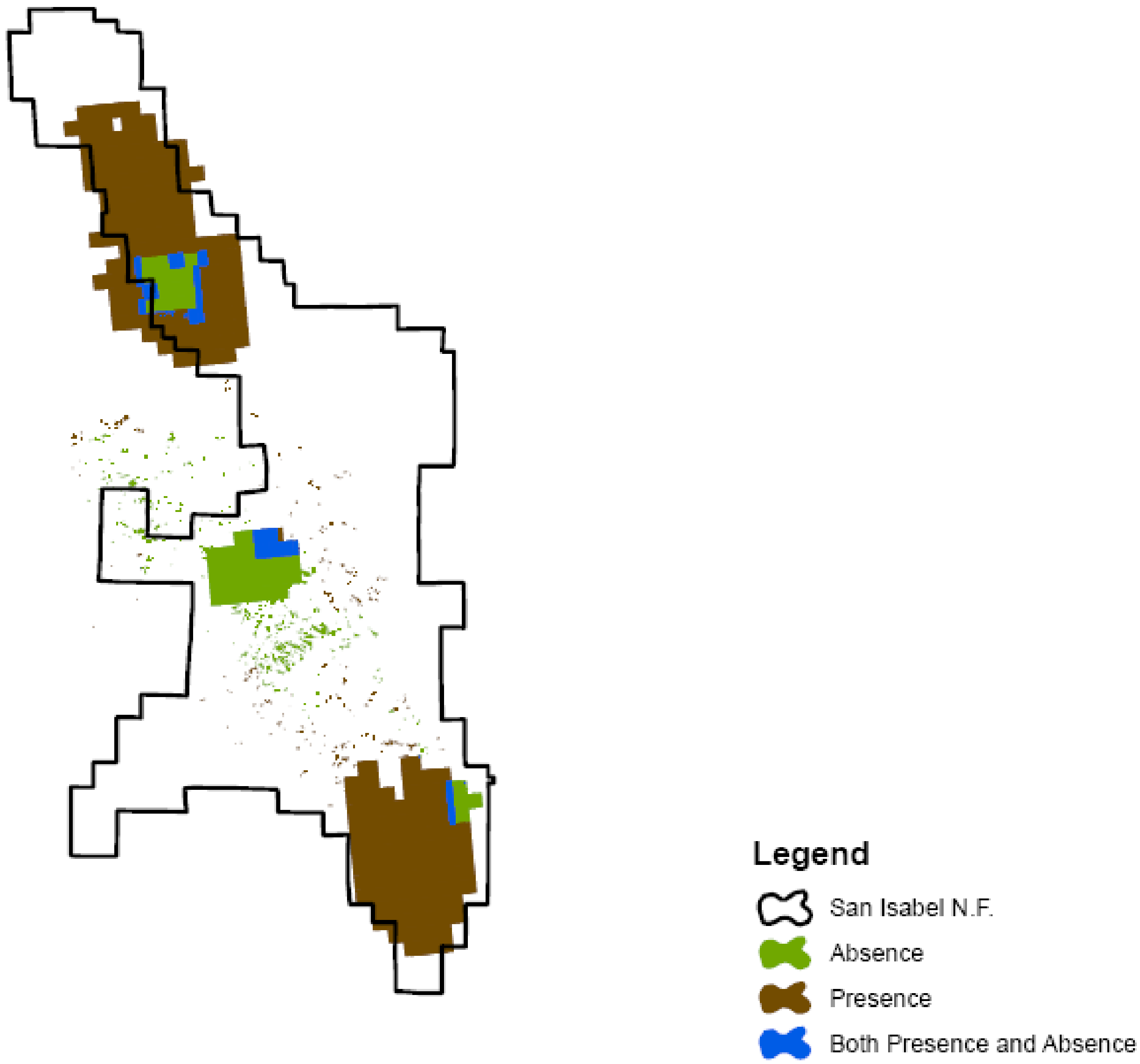


Figure 2. The WPBR presence-absence model as run by Kearns (2005), using the Wet Mountains on the San Isabel National Forest as an example. Note the overlapping polygons with contradictory results, indicating both a presence and an absence of WPBR in the same location.

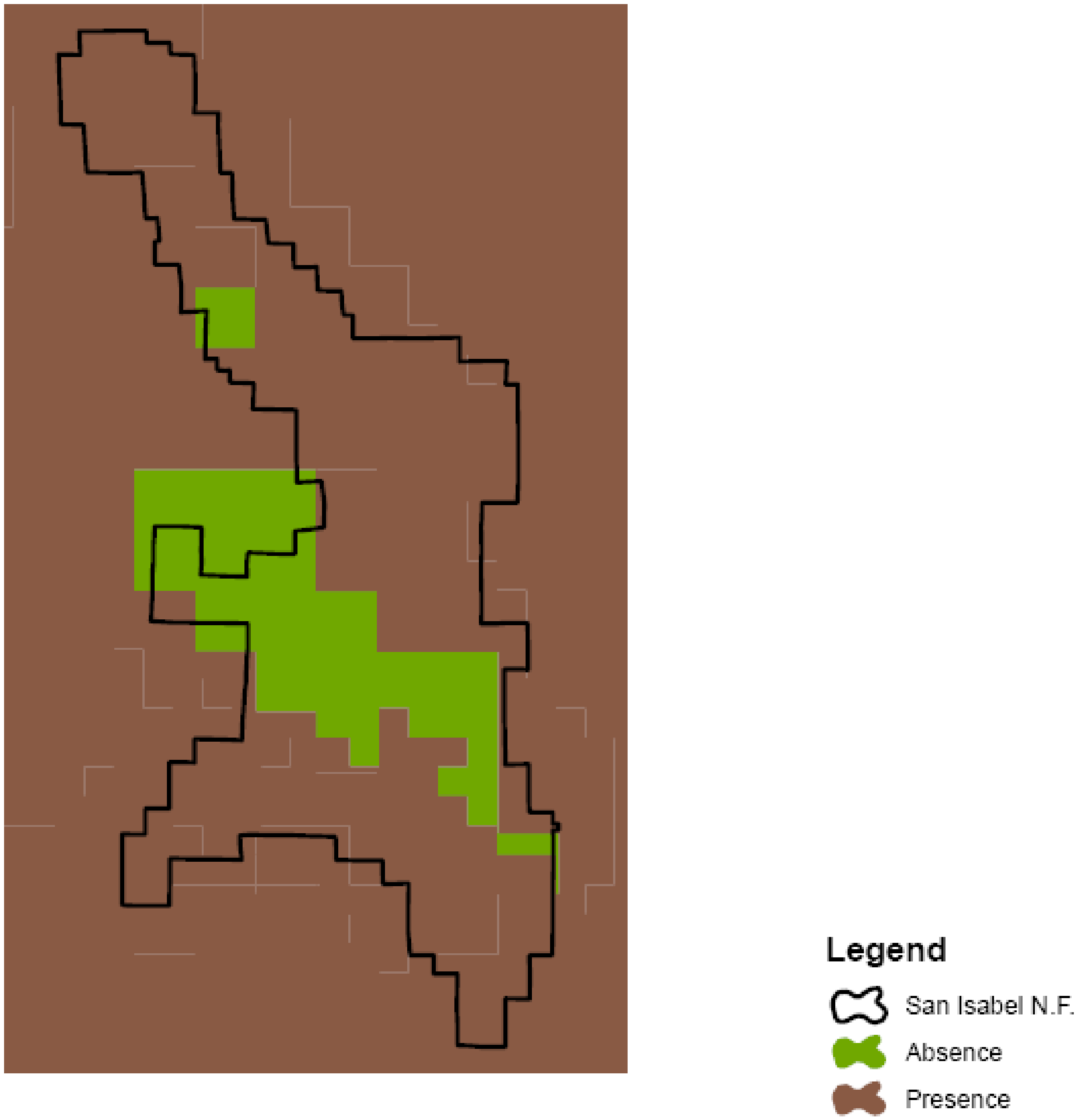


Figure 3. The output of the Kearns (2005) WPBR presence-absence model after it was re-run based on weather variables from the PRISM dataset for the Wet Mountains on the San Isabel National Forest. Much of the area covered does not support populations of white pine.

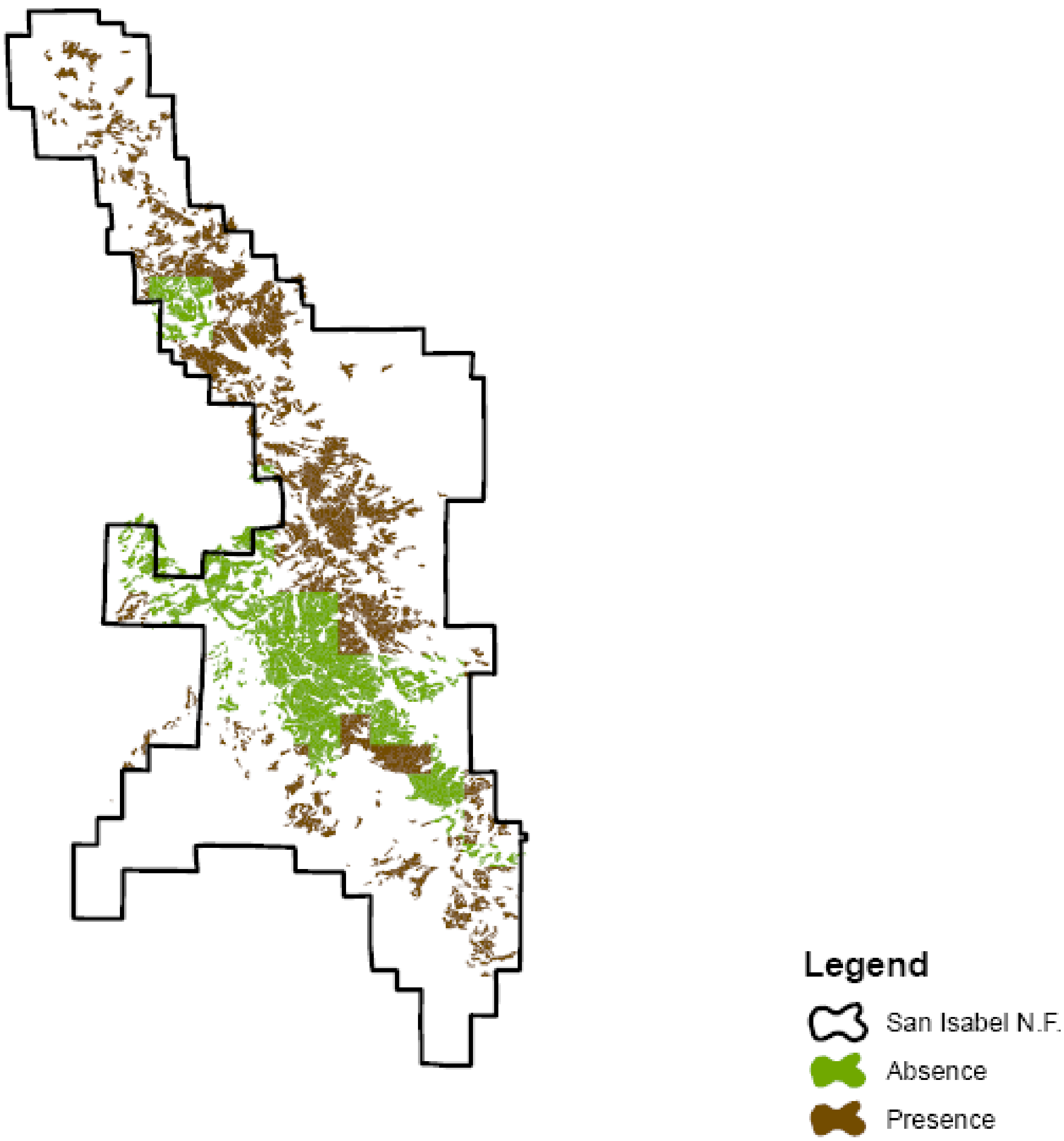


Figure 4. The output of the Kearns (2005) WPBR presence-absence model, after it was re-run based on weather variables from the PRISM dataset, has been spatially joined with polygons containing a white pine component from the preferred vegetation coverage for the Wet Mountains on the San Isabel National Forest, the Wet Mountain Common Vegetation Unit (CVU) dataset.

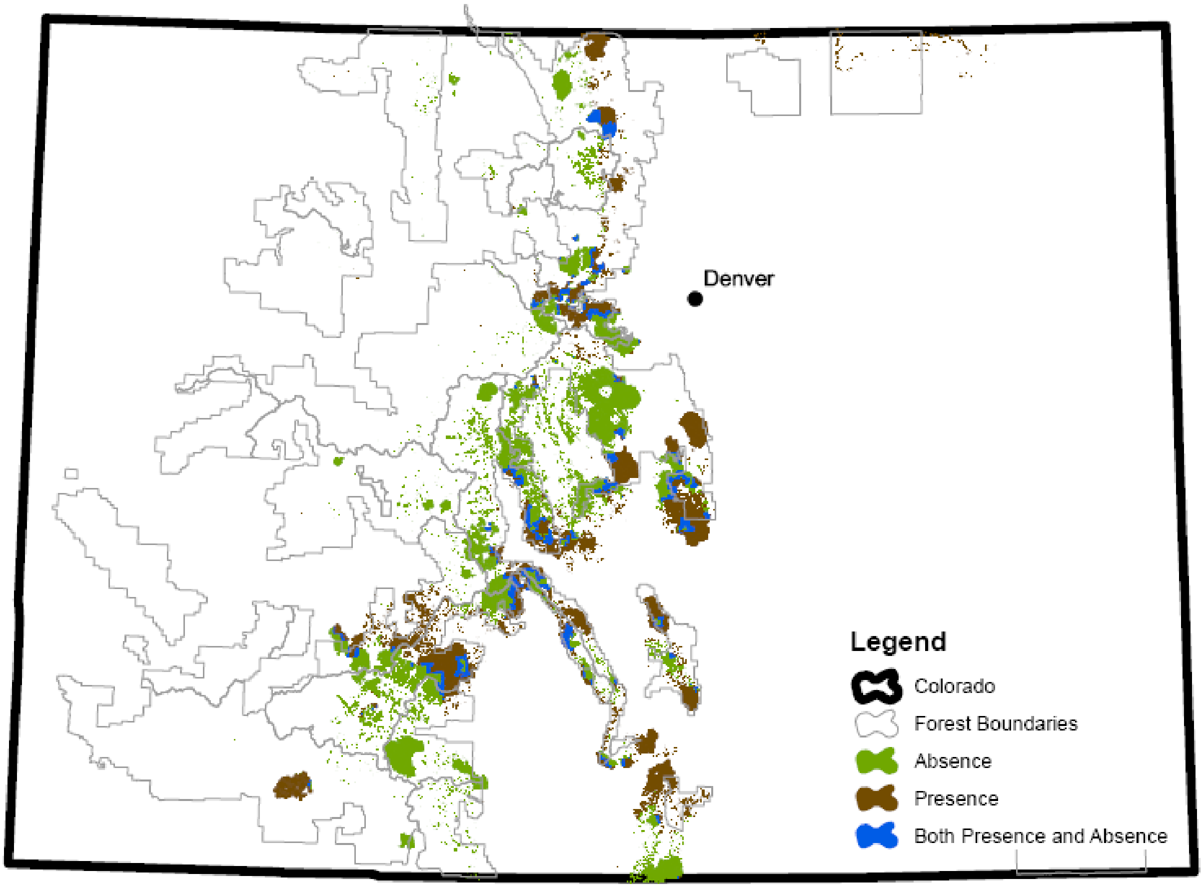


Figure 5. The WPBR presence-absence model as run by Kearns (2005) for all of Colorado. Note the overlapping polygons with contradictory results, indicating both a presence and an absence of WPBR in the same location.

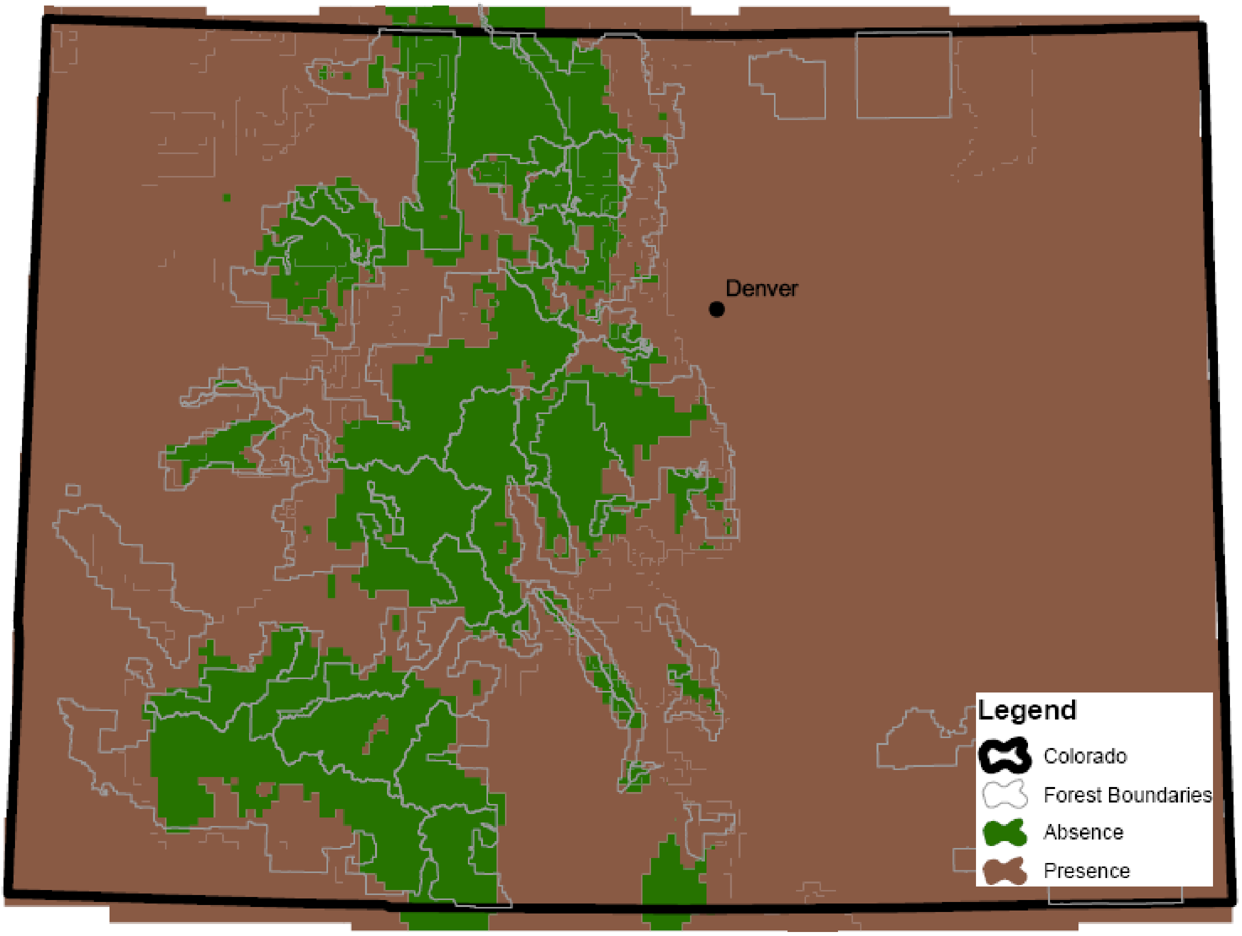


Figure 6. The output of the Kearns (2005) WPBR presence-absence model after it was re-run based on weather variables for the entire state of Colorado. The model output can be joined with white pine vegetation coverages and used to evaluate risk in areas where white pine is known to be present.

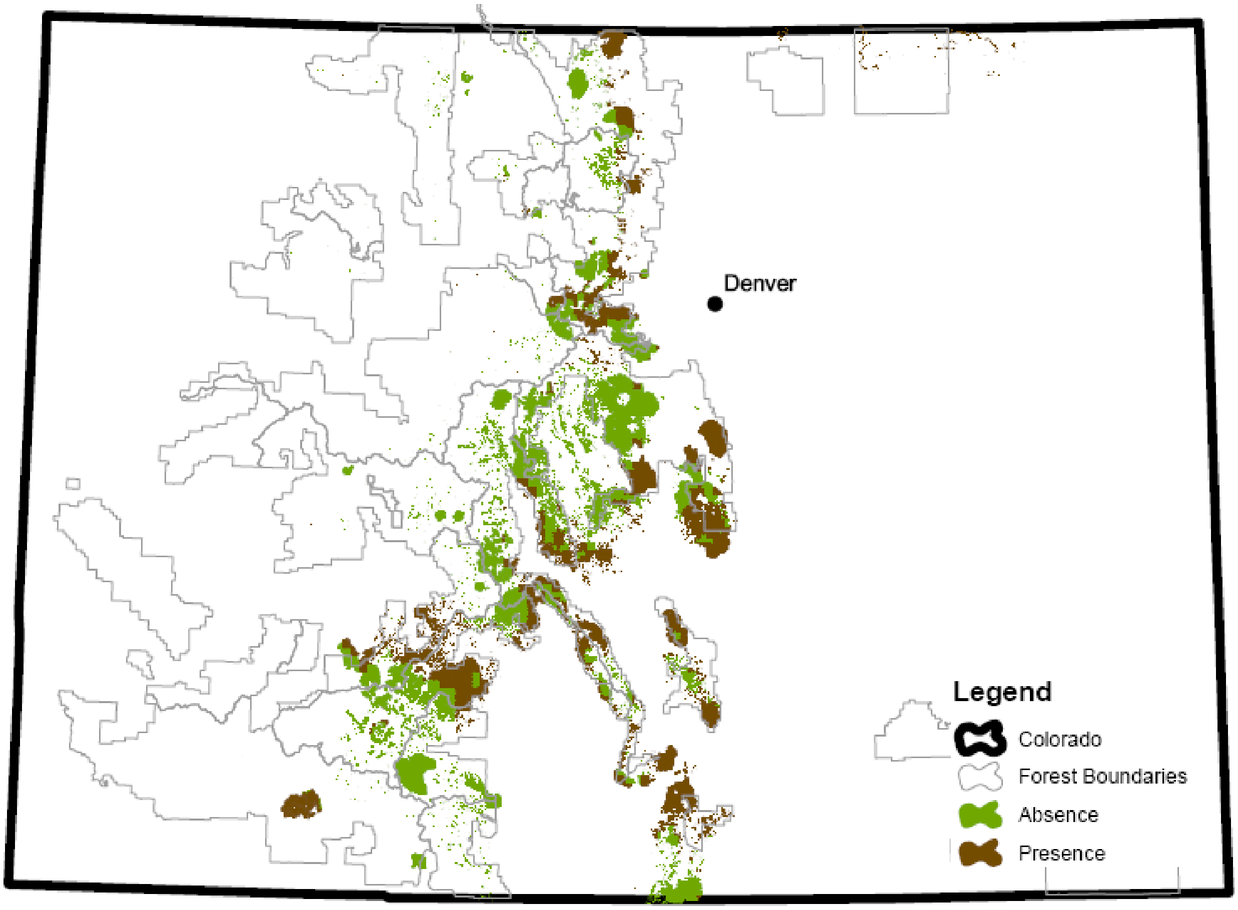


Figure 7. The output of the Kearns (2005) WPBR presence-absence model after it was re-run based on weather variables for the entire state of Colorado. Here, the model output was spatially joined with the combination of white pine polygons from the four vegetation coverages originally used by Kearns (2005).

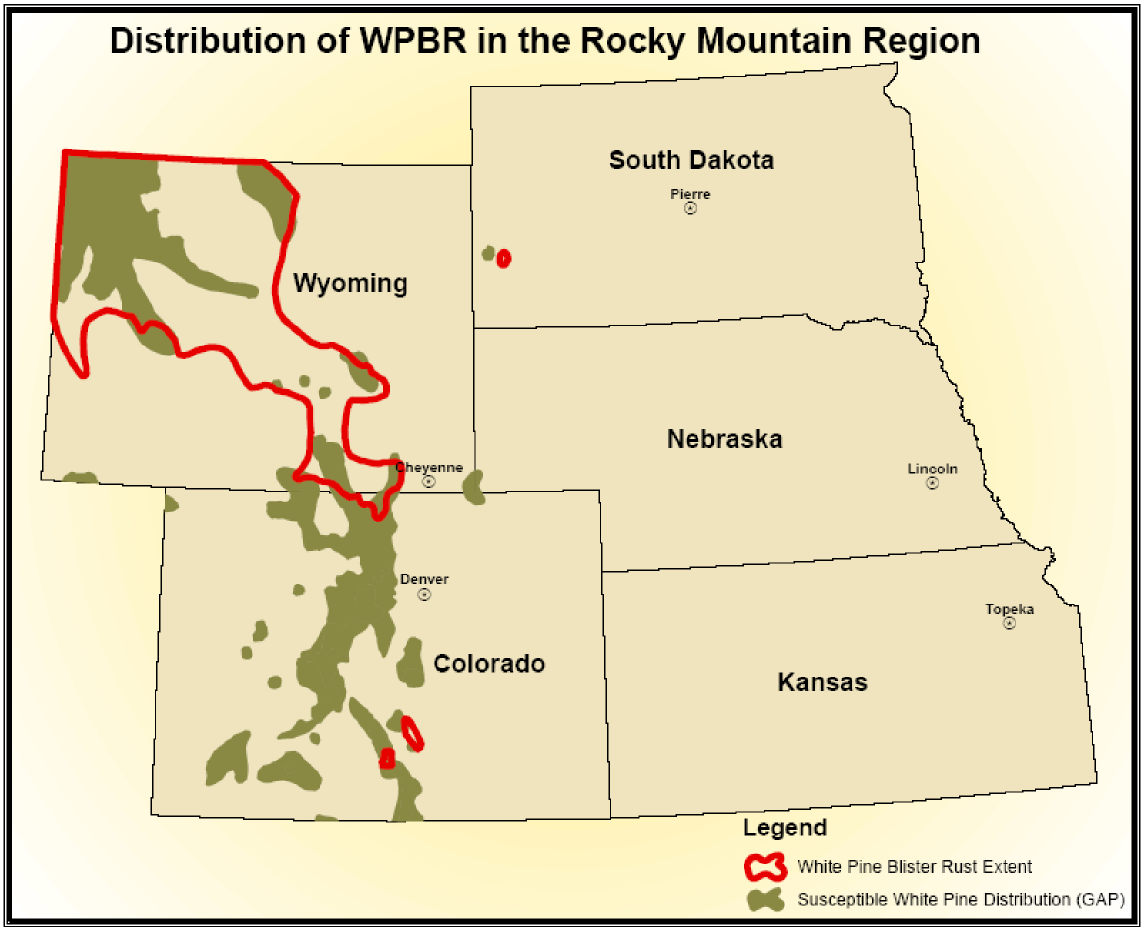


Figure 8. The known extent of WPBR in the Rocky Mountain Region (USDA Forest Service Region 2) as of January, 2006.