# Methods

Forward-stepping discriminant analysis was conducted on a random sample of n = 5305 (or 80% of the data) observations of cheatgrass cover. The model was validated with the remaining n = 1325 observations. Cover was classified as Low for cover ≤ 2% and High for cover > 2% based on a clear cut-off value in the distribution as visualized by a histogram of the cover for the model building data (Figure 1). Initially 50 variables were considered for model-building including elevation, potential radiation index, two NDVI based variables, cumulative precipitation (Oct-April of the following year), cumulative growing degree days, and 30-yr average monthly minimum and maximum temperature and monthly and annual precipitation. An average minimum winter temperature (Nov-Dec), average maximum winter (Nov-Feb) temperature, Spring (Apr-May), Summer (Jun-Sep), and Winter (Dec-Feb) precipitation were calculated. One of each pair of variables was deleted from model-building when the pair-wise correlation exceeded r = 0.8. Variables selected for model-building (n = 19) were less correlated, had smaller interquartile ranges for the Low cover class among competing variables, and greater separation between the Low and High cover class quartiles. Variables were standardized with the mean and standard deviation of the model data before discriminant analysis. To increase the separation between cover classes, observations with cover greater than 2% but less than or equal to 10% were excluded from model development. However, the discriminant score and predicted classification was calculated for all observations. A confusion matrix was calculated for both the model and validation data and was used to evaluate the resulting model.



Figure 1. Histogram of cheatgrass cover (%) for a random sample of n = 5305 used for model building. Bins are each equal to 2%.

# Steps for classifying image based on the discriminant functions:

## Discrimination

**Step 1.** Calculate the variables needed for **the model**:

* elev
* prr
* zmaxyr\_dnd
* ltmed\_pndv
* ltmed\_sumg
* tmin03
* tmin11
* tmax05
* W-avgMaxT = average(tmax11, tmax12, tmax01, tmax02)
* ppt03
* ppt06
* ppt07
* Winter\_PPT = Sum(ppt12, ppt01, ppt02)

**Step 2.** Standardize each variable by subtracting the mean and then dividing by the standard deviation = (x-mean)/stdev using the tabled means and standard deviations below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | elev | prr | zmaxyr\_dnd | ltmed\_pndv | ltmed\_sumg |
| Mean | 6262 | 10527 | 198 | 6706 | 779 |
| Stdev | 1818 | 2594 | 390 | 469 | 212 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | tmin03 | tmin11 | tmax05 | W-avgMaxT | ppt03 | ppt06 | ppt07 | Winter\_PPT |
| Mean | -3.55 | -4.30 | 19.2 | 4.78 | 28.8 | 23.6 | 14.6 | 85.9 |
| Stdev | 2.01 | 1.98 | 2.19 | 2.04 | 10.7 | 9.16 | 8.99 | 37.1 |

**Step 3.** Calculate score for **Model**:

Score1 = 0.348\*elev + 0.115\*prr + 0.110\*zmaxyr\_dnd + 0.327\*ltmed\_pndv + 0.430\*ltmed\_sumg +

(-0.094)\*tmin03 + 0.429\*tmin11 + 0.676\*tmax05 + (-0.380)\*W-avgMaxT + (-0.450)\*ppt03 +

0.201\*ppt06 + (-0.637)\*ppt07 + 0.519\*W\_PPT + 0.059

|  |  |
| --- | --- |
| Raw Coefficients (Stand Data Model.sta) for Canonical Variables Exclude condition: v4="Moderate" | |
| Variable | Root 1 |
| elev | 0.348 |
| prr | 0.115 |
| zmaxyr\_dnd | 0.110 |
| ltmed\_pndv | 0.327 |
| ltmed\_sumg | 0.430 |
| tmin03 | -0.094 |
| tmin11 | 0.429 |
| tmax05 | 0.676 |
| W-avgMaxT | -0.380 |
| ppt03 | -0.450 |
| ppt06 | 0.201 |
| ppt07 | -0.637 |
| Winter\_PPT | 0.519 |
| Constant | 0.059 |
| Eigenval | 0.358 |
| Cum.Prop | 1.000 |

**Step 4.** Calculate distance from centroids for **Model**:

Low: C <= 2% and High: C > 2%

Distance from Low = sqrt[ (Score1 + 0.501)2 ]

Distance from High = sqrt[ (Score1 – 0.714)2 ]

|  |  |
| --- | --- |
| Group | Means of Canonical Variables (Stand Data Model.sta) Exclude condition: v4="Moderate" |
| Root 1 |
| Low | -0.501 |
| High | 0.714 |

**Step 5.** Calculate the probability of class membership:

SDist = Sum of the two distances from Step 4

P-Low = Probability in Low = 1 – Distance from Low/SDist

P-High = Probability in High = 1 – Distance from High/SDist

**Step 6.** Estimated Class

1. If P-High >P-Low, then assign “High”
2. ELSE “Low”

**Model Data Confusion Matrix**

|  |  |  |  |
| --- | --- | --- | --- |
| Estimated | Observed | | P-Correct |
| Low | C>2% |
| Low | 1916 | 822 | 74% |
| C>2% | 686 | 1881 | 70% |
| Total | 2602 | 2703 | 72% |
| % in Error | 26% | 30% |  |

**Verification Data Confusion Matrix**

|  |  |  |  |
| --- | --- | --- | --- |
| Estimated | Observed | | P-Correct |
| Low | C>2% |
| Low | 454 | 231 | 74% |
| High | 157 | 503 | 69% |
| Total | 611 | 734 | 71% |

**This is corrected for analysis without sampYr\_DND.**

**Without sampyr\_DND**

Number of variables in the model: 13

Wilks' Lambda: .7363446 approx. F (13,4415) = 121.6026 p < 0.0000

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| N=4429 | Discriminant Function Analysis Summary (Stand Data Model.sta) No. of vars in model: 13; Grouping: Cover Class (2/10) (2 grps) Wilks' Lambda: .73634 approx. F (13,4415)=121.60 p | | | | | |
| |  | | --- | | Wilks' Lambda | | |  | | --- | | Partial Lambda | | |  | | --- | | F-remove 1,4415 | | |  | | --- | | p-value | | |  | | --- | | Toler. | | |  | | --- | | 1-Toler. (R-Sqr.) | |
| |  | | --- | | elev | | 0.739340 | 0.995949 | 17.9587 | 0.000023 | 0.144095 | 0.855906 |
| |  | | --- | | prr | | 0.738686 | 0.996830 | 14.0382 | 0.000181 | 0.956253 | 0.043747 |
| |  | | --- | | zmaxyr\_dnd | | 0.738575 | 0.996980 | 13.3725 | 0.000258 | 0.940701 | 0.059299 |
| |  | | --- | | ltmed\_pndv | | 0.749713 | 0.982168 | 80.1563 | 0.000000 | 0.607192 | 0.392808 |
| |  | | --- | | ltmed\_sumg | | 0.739569 | 0.995640 | 19.3329 | 0.000011 | 0.102429 | 0.897571 |
| |  | | --- | | D\_tmin03 | | 0.736392 | 0.999936 | 0.2837 | 0.594326 | 0.032186 | 0.967814 |
| |  | | --- | | D\_tmin11 | | 0.737745 | 0.998101 | 8.3988 | 0.003773 | 0.044403 | 0.955597 |
| |  | | --- | | D\_tmax05 | | 0.748390 | 0.983905 | 72.2217 | 0.000000 | 0.141204 | 0.858796 |
| |  | | --- | | D\_W-avgMaxT | | 0.742726 | 0.991408 | 38.2621 | 0.000000 | 0.228150 | 0.771850 |
| |  | | --- | | D\_ppt03 | | 0.739953 | 0.995123 | 21.6360 | 0.000003 | 0.091209 | 0.908792 |
| |  | | --- | | D\_ppt06 | | 0.738565 | 0.996994 | 13.3125 | 0.000267 | 0.275188 | 0.724812 |
| |  | | --- | | D\_ppt07 | | 0.761418 | 0.967070 | 150.3365 | 0.000000 | 0.326746 | 0.673254 |
| |  | | --- | | D\_Winter PPT | | 0.741514 | 0.993029 | 30.9943 | 0.000000 | 0.099495 | 0.900505 |

|  |  |  |  |
| --- | --- | --- | --- |
| Group | Classification Matrix (Stand Data Model.sta) Rows: Observed classifications Columns: Predicted classifications Exclude condition: v4="Moderate" | | |
| |  | | --- | | Percent Correct | | |  | | --- | | Low p=.50000 | | |  | | --- | | High p=.50000 | |
| |  | | --- | | Low | | 73.63567 | 1916 | 686 |
| |  | | --- | | High | | 73.94636 | 476 | 1351 |
| |  | | --- | | Total | | 73.76383 | 2392 | 2037 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Roots Removed | Chi-Square Tests with Successive Roots Removed (Stand Data Model.sta) Exclude condition: v4="Moderate" | | | | | |
| |  | | --- | | Eigen- value | | |  | | --- | | Canonicl R | | |  | | --- | | Wilks' Lambda | | |  | | --- | | Chi-Sqr. | | |  | | --- | | df | | |  | | --- | | p-value | |
| |  | | --- | | 0 | | 0.358060 | 0.513474 | 0.736345 | 1352.926 | 13 | 0.00 |

|  |  |
| --- | --- |
| Variable | Raw Coefficients (Stand Data Model.sta) for Canonical Variables Exclude condition: v4="Moderate" |
| |  | | --- | | Root 1 | |
| |  | | --- | | elev | | 0.348104 |
| |  | | --- | | prr | | 0.114603 |
| |  | | --- | | zmaxyr\_dnd | | 0.109789 |
| |  | | --- | | ltmed\_pndv | | 0.327361 |
| |  | | --- | | ltmed\_sumg | | 0.429661 |
| |  | | --- | | D\_tmin03 | | -0.094050 |
| |  | | --- | | D\_tmin11 | | 0.428860 |
| |  | | --- | | D\_tmax05 | | 0.676334 |
| |  | | --- | | D\_W-avgMaxT | | -0.380012 |
| |  | | --- | | D\_ppt03 | | -0.449677 |
| |  | | --- | | D\_ppt06 | | 0.200735 |
| |  | | --- | | D\_ppt07 | | -0.636680 |
| |  | | --- | | D\_Winter PPT | | 0.519449 |
| |  | | --- | | Constant | | 0.059313 |
| |  | | --- | | Eigenval | | 0.358060 |
| |  | | --- | | Cum.Prop | | 1.000000 |

|  |  |
| --- | --- |
| Group | Means of Canonical Variables (Stand Data Model.sta) Exclude condition: v4="Moderate" |
| |  | | --- | | Root 1 | |
| |  | | --- | | Low | | -0.501297 |
| |  | | --- | | High | | 0.713944 |