USING LANDSCAPE CHARACTERISTICS AS PRIOR INFORMATION FOR BAYESIAN CLASSIFICATION OF REMOTELY SENSED IMAGERY

William J. Price and Bahman Shafii
Statistical Programs
Lawrence W. Lass and Timothy S. Prather
Division of Plant Sciences

College of Agricultural and Life Sciences
University of Idaho
Moscow, ID 83844

Yellow starthistle is a dominant weed of north-central Idaho canyon grasslands. The distribution of yellow starthistle can be affected by general landscape characteristics, such as land use, as well as specific terrain related features such as elevation, slope, and aspect. Slope and aspect can be considered as indicators of plant community composition and distribution. Hence, these variables may be incorporated into prediction models to estimate the likelihood of yellow starthistle occurrence. An empirically derived nonlinear model based on landscape characteristics was developed to predict the likelihood of yellow starthistle occurrence in north central Idaho (Shafii, et al. 1999). While the model was employed to predict the invasion potential of yellow starthistle into new areas, it could also be used as auxiliary data for classifying this weed species in remotely sensed imagery. To accomplish this, the predicted values of the model are regarded as prior information on the presence of yellow starthistle. A Bayesian image classification algorithm using this prior information is then applied to a corresponding set of remotely sensed data. The end result is a map indicating the posterior probabilities of yellow starthistle occurrence given the landscape characteristics. This technique is demonstrated considering the presence and absence of prior information and is shown to result in lower omissional and commissional error rates when the landscape characteristics are utilized.

Keywords: nonlinear model, image classification, posterior probabilities, error rates

I. INTRODUCTION

Yellow starthistle (*Centaurea solstitialis* L.) is an introduced toxic weed which infests over 15 million acres in the western United States. It can cause serious economic losses due to its toxicity to horses, potential for crop yield reduction, and reduction of native plant diversity. Major invasions have occurred on rangeland and non-crop lands, however, cultivated lands such as dryland grain, legume, seed crop, and pasture are also susceptible to invasion (Lass, et al. 1999). Yellow starthistle thrives best on warm, well-drained soils with 15-30 inches of annual precipitation. Infestations are common in the canyon grasslands of Northern Idaho, California, Oregon, and Washington and have become established in many Western States.

Early detection, as well as the ability to predict invasion potential are important

considerations for the strategic management of yellow starthistle. While classification of remotely sensed imagery can be a useful tool in identifying weed infestations, traditional image classification techniques can result in high omissional error rates when detecting new yellow starthistle infestations. In addition, land managers must focus their detection efforts on areas with high likelihoods of invasion to best utilize their limited resources. The objective of this paper is to improve the classification accuracy of yellow starthistle as well as identify areas at risk for infestation. The methods used incorporate information on landscape variables and their relationship to yellow starthistle occurrence into the process of image classification.

II. MATERIALS AND METHODS

Classification. The data to be classified are composed of p images, each representing a specific spectral band. Individual pixels within an image record the intensity of the light reflected from the ground. One means of classifying such data is the use of a multivariate quadratic discriminant function (see for example Theodoridis and Kouttroumbas, 1999). Classification begins by defining the likelihood of the Tth class as:

$$L_T = (2\pi)^{-P/2} \, |V_T|^{-1/2} \, \exp(-.5M_T) \, , \tag{1}$$

where V_T is the variance-covariance matrix for the p spectral bands in an independent training data set for class T, and M_T is the Manhalanobis distance given by:

$$M_T = (d - \overline{x_T})' \, V_T^{-1} \, (d - \overline{x_T}) \tag{2}$$

Equation (2) measures the weighted distance between an unknown data vector, d, and the corresponding average vector for the Tth class in the training data. Once the likelihood is formed, the posterior probability of belonging to class T is given by:

$$p(T|d) = \frac{q_T \, L_T}{\sum q_T \, L_T} \tag{3}$$

where q_T is the prior probability of belonging to class T given d and the denominator is a normalization factor that is summed over all classes.

Traditionally, the prior is either specified as uniform across all classes, or is set proportional to the size of each class in the training data. While these priors provide reference points for classification assessment, in practice, they result in constant prior values within each class and fail to take advantage of the spatial structure that may be present in the remotely sensed imagery. As an alternative, prior values may be specified according to spatial characteristics associated with defined classes.

Spatial Prior. For Northern Idaho, a relationship between yellow starthistle occurrence and landscape variables has previously been shown (Shafii et al, 1999). Utilizing USGS digital

elevation model (DEM) data and a University of Idaho yellow starthistle survey, the proportion of yellow starthistle present was determined within slope and aspect classes defined as :
Aspect: A_i ; every 22.5 ° ; i = 1, 2, 3, ..., 16, and
Slope: S_j ; 0-10%, every 5% from 10% to 60% and > 60% ; j = 1, 2, 3, ..., 16.
A polar coordinate transformation non-linear model of the form:

$$Z_{ij} = b_0 + b_1X + b_2X^2 + b_3Y + b_4Y^2 + b_5\ln(X^2 + Y^2) + e_{ij} \qquad (4)$$

was then constructed where
$\qquad X = S_j*\cos(A_i),$
$\qquad Y = S_j*\sin(A_i),$
$\qquad Z_{ij} =$ a logit transformation of the yellow starthistle proportions,
$\qquad b_0 - b_5 =$ estimated parameters, and
$\qquad e_{ij} =$ an error term under the usual assumptions of regression.
The estimated logits from equation (4) may be untransformed as $q_T = \exp(Z_{ij})/(1 + \exp(Z_{ij}))$ and incorporated into the classification algorithm given above. The result is a prior probability image where the values vary according to the slope and aspect (spatial position) of each pixel.

III. EMPIRICAL RESULTS

The study area included parts of the Salmon River in North-Central Idaho and was composed primarily of mixed rangeland (81%) with smaller proportions of pasture (0.5%) and evergreen forest (18.5%). Images for this area were recorded in July 2000 using a Probe 1 hyperspectral sensor [Earth Search Sciences Inc., Kalispell, MT] at a resolution of 5 x 5 m for each pixel. The resulting data included 128 spectral bands in the visible to near infrared range (450-2500 nm). Due to computational limitations, not all of these bands could be used. Based on previous experience with yellow starthistle, seven bands were selected: 520, 600, 680, 770, 890, 1020, and 1150 nm. These values covered the available range of data including regions known to be active in photosynthesis. The necessary data for computing prior values, q_T, were obtained from 7.5 minute USGS DEM data and a University of Idaho yellow starthistle survey. All data sources were collated to a common grid representing 16.3 million 5m x 5m cells.

Five identifiable classes were designated for the purpose of classification. These were annual grasses (AGRASS), yellow starthistle (YST), perennial grasses (GRASS), River (RIVER), and evergreen forest (FOREST). The last two categories were combined for post classification accuracy assessment.

All computations and graphics were carried out using SAS (1991) or custom C programs. Program codes are available from the authors at http://www.uidaho.edu/ag/statprog.

Figure 1 shows a sample spectral image of the study area at 890 nm. The inset image is a higher resolution enlargement of a relevant section. The Salmon River appears as a dark band running across the image. Above the river to the north are south facing canyon slopes where yellow starthistle and annual grasses can be expected to appear. South of the river are north facing slopes which are typically covered with forest and perennial grasses. The highlighted areas represent verification sites which are known to have high densities of yellow starthistle and

were used for accuracy assessment. For clarity of presentation, the subsequent figures below will concentrate only on these verification areas. A more detailed set of colored images for a larger area can be found at the URL listed above.

Uniform prior.

Initially, the prior probabilities were assumed to be uniform, i.e. $q_T = 0.2$. A portion of the resulting posterior probability map for yellow starthistle is shown in Figure 2. Here, probabilities are coded on a gray scale with black being 0.0 and white being 1.0. Although the pixels that were classified as YST were high in probability (the colors were bright), the uniform prior resulted in large sections of the verification sites showing low or zero probability (dark or black color). The training sites, on the other hand, were almost 100% accurately identified (not shown). If probability levels greater than or equal to 0.3 can be assumed to be YST, an error matrix for the image can be calculated as shown in Table 1. In this type of cross classification table, the diagonal values represent correctly classified pixels. Values off the diagonal represent mis-classifications which are summarized by the proportions given on the table margins. For example, the proportion of pixels committed to any given class is called the commissional error, $\hat{C}_i$, and is shown on the right hand margin. For YST, 1% of the true yellow starthistle pixels were committed to other classes. Similarly, the proportion of pixels omitted from a class or the omissional error, $\hat{O}_i$, is given at the bottom of the table. In this case yellow starthistle has an omissional error rate of 59%, which is quite high (for more information on error matrices and error rates in remote sensing, see for example Congalton, 1991 or Shafii and Price, 2001). An examination of values within Table 1 shows that a large number of pixels, 5429, were omitted from YST and placed in the AGRASS category. This was not necessarily unexpected since the two species share similar habitats. When the probability maps for the two categories were compared, the low probability areas for YST appeared as high probability areas for AGRASS. This would indicate some confusion between the classes for the classification algorithm, reason for which can be illustrated by examining the spectral responses of each class. In Figure 3, box plots of the spectral responses for 890 nm are given for four categories: the training sites for YST and AGRASS, the correctly classified YST pixels, and the pixels omitted from YST into AGRASS. From this plot, it is obvious that the correctly classified YST pixels matched the YST training data closely while the omitted pixels were more similar to the AGRASS training data. Similar patterns could be seen for the other spectra and, in fact, were evident across all 128 spectral images. This is most likely due to a strong spectral response from AGRASS and a combination of the two species within a pixel. If the objective is to identify YST pixels regardless of mixing, then a procedure to filter the AGRASS pixels for the presence of YST is needed. The previously developed YST prediction model (Shafii et al, 1999) can be used for this purpose.

Spatial prior.

The model shown in equation (4) was estimated and validated for three land use categories: predominately pasture, mixed rangeland, and forest (Shafii et al, 1999). An example of the observed and predicted surfaces for mixed rangeland is shown in Figure 4. The model indicates that the expected proportion of yellow starthistle systematically varied with changes in slope and aspect. Using these predicted proportions, a prior probability map was developed that provided values of q_T for each pixel based on slope and aspect landscape variables. These prior

values were then assigned to both the YST and AGRASS categories. The remaining probability was then split uniformly among the other categories. The resulting posterior probability classification for YST is shown in Figure 5. While probability levels are not always maximized, the verification sites now appear to be correctly filled in with reasonable values. Thus, the application of the prior information regarding landscape variables did improve the classification.

Unfortunately for most applications, probability maps are not directly useful. Users of classification maps typically require discrete or "hardened" categories that indicate strict presence or absence, not statements of probability. In order to create a discrete map, a cut-off or threshold probability must be defined as was done above for the uniform prior. However, these cut-off values will vary according to the needs of the users. Furthermore, any resulting omissional or commissional error rates will be sensitive to this choice. To demonstrate this, several threshold values ranging from 0.0 to 0.95 were assigned to the YST class for both the uniform and spatial prior classifications. At each threshold, the omissional and commissional error rates were computed. The results are displayed in Figures 6 and 7, respectively. For both priors, increasing threshold values lead to higher omissional error rates. This was expected as higher thresholds imply that a user must be more certain about the classification and, therefore, will omit more pixels. In all cases, the spatial prior out performed the uniform prior with consistently lower omissional error rates. If the spatial prior was used, a reasonable cut-off value might be between 0.3 and 0.4 where the error rate began to increase. The commissional error gave an opposite trend with error rates decreasing as the threshold increased. While the spatial prior appeared to have error rates higher than the uniform prior, the overall effect was negligible since commissional error rates never exceeded 3% at any threshold level.

Using the 0.3 probability threshold, an error matrix for the spatial prior was computed (Table 2). The spatial prior produced a decrease in the omissional error rate over that of the uniform prior of more than 50%. The number of pixels omitted from YST into AGRASS was reduced from 5429 to 1280 and correctly classified YST more than doubled from 4997 to 9040 pixels. Commissional error rates were essentially unchanged. Although the commissional and omissional errors for AGRASS were high, they were inconsequential in this case because the objective was to identify areas with YST and mis-classification of the other categories was not of any concern.

IV. CONCLUDING REMARKS

Identification of sparse infestations is an important consideration for the management and control of yellow starthistle. Traditional remote sensing methods provide appropriate tools for locating weed infestations, but they may fail to identify areas of low density yellow starthistle. By using prior knowledge concerning the relationship between yellow starthistle and landscape variables such as slope and aspect, the ability to classify yellow starthistle in remotely sensed imagery is greatly improved. Separation of classes with similar spectral signatures is achieved and omissional error rates are reduced. The resulting probability maps, however, must be "hardened" into discrete classes for use in field situations. This process must be monitored for its effects on both omissional and commissional error rates.

V. REFERENCES

Congalton, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing Environ. 37 pp. 35-46.

Lass, L. W. J. P. McCaffrey, D. C. Thill and R. H. Callihan. 1999. Yellow starthistle biology and management in pasture and rangeland. University of Idaho Bulletin No. 805, 18 pp.

SAS Institute Inc. 1991. SAS/STAT User's Guide, Version 6 Fourth Ed., Vol. 2. Sas Institute Inc, Cary NC.

Shafii, Bahman and W. J. Price. 2001. Application of Bayesian Methods for Assessing the Detection Accuracy in Remote Sensing. American Statistical Association, Section on Bayesian Statistics, CD-ROM.

Shafii, B., W. J. Price, L. W. Lass and D. C. Thill. 1999. Estimating the Likelihood of Yellow Starthistle Occurrence Using an Empirically Derived Nonlinear Regression Model. Proceedings of the 1994 Kansas State University Conference on Applied Statistics in Agriculture. Kansas State University, Manhattan, Ka.

Theodoridis, S. and Koutroumbas, K. 1999. Pattern Recognition. Academic Press, NY, 625 pp.

Uniform prior

		Ground Truth				
		AGRASS	YST	GRASS	OTHER	$\hat{C}_i$
Classified	AGRASS	407	5429	0	0	0.93
	YST	45	4997	0	0	0.01
	GRASS	0	0	1183	32	0.03
	OTHER	181	1741	473	6269	0.28
	$\hat{O}_i$	0.36	0.59	0.29	0.01	

Table 1. Error matrix for the uniform prior classification.

Spatial prior	Ground Truth				
	AGRASS	**YST**	**GRASS**	**OTHER**	**$\hat{C}_i$**
AGRASS	191	1280	0	0	0.87
YST	219	9040	0	0	0.02
GRASS	0	0	1183	32	0.03
OTHER	223	1847	473	6269	0.29
$\hat{O}_i$	0.70	0.28	0.29	0.01	

Classified (row label, left side)

Table 2. Error matrix for the spatial prior classification.
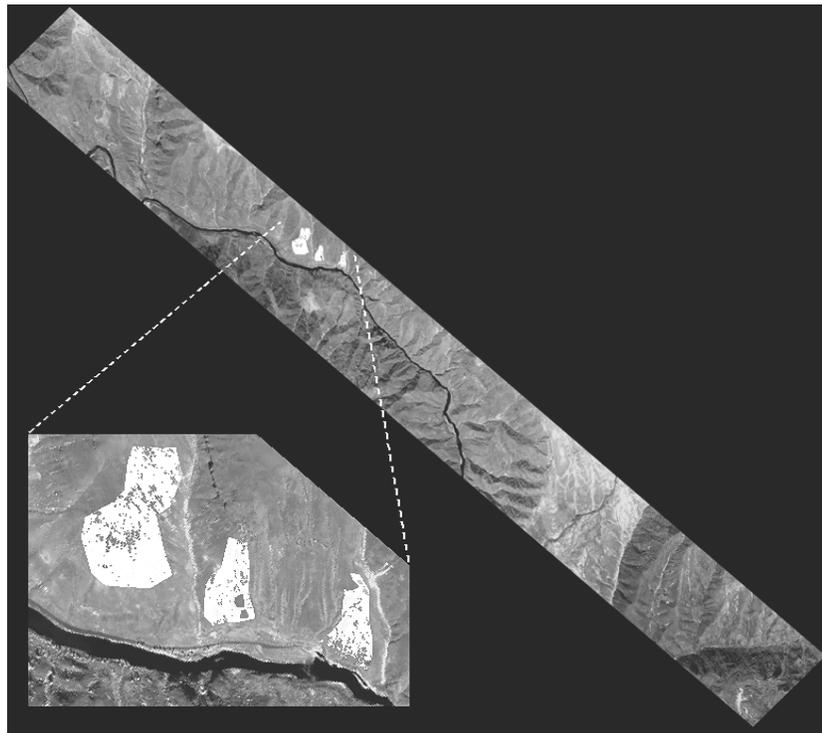
Figure 1. Yellow starthistle study area along the Salmon River in Northern Idaho. The highlighted areas indicate verification sites.
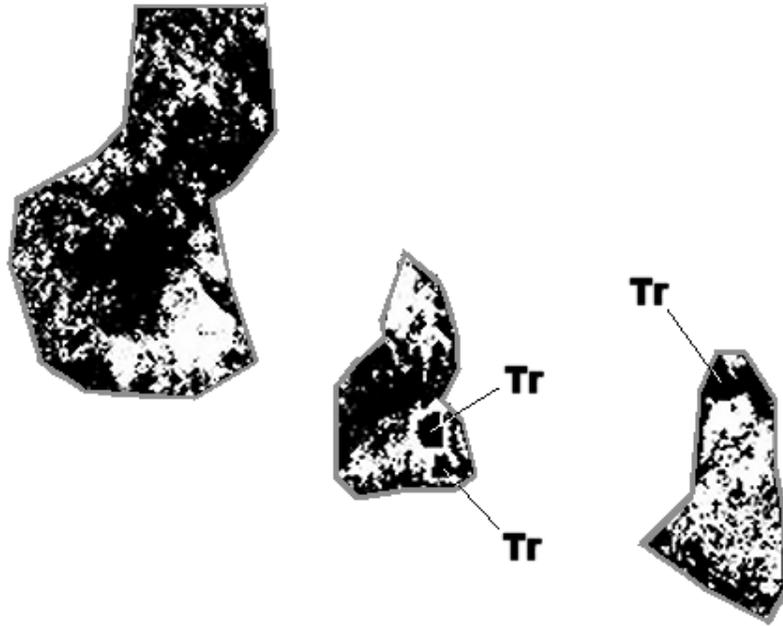
Figure 2. Probability map for yellow starthistle assuming a uniform prior. The probability levels are indicated on a gray scale with black = 0.0 and white = 1.0. Areas marked as **Tr** were training sites and were not included in accuracy assessment.
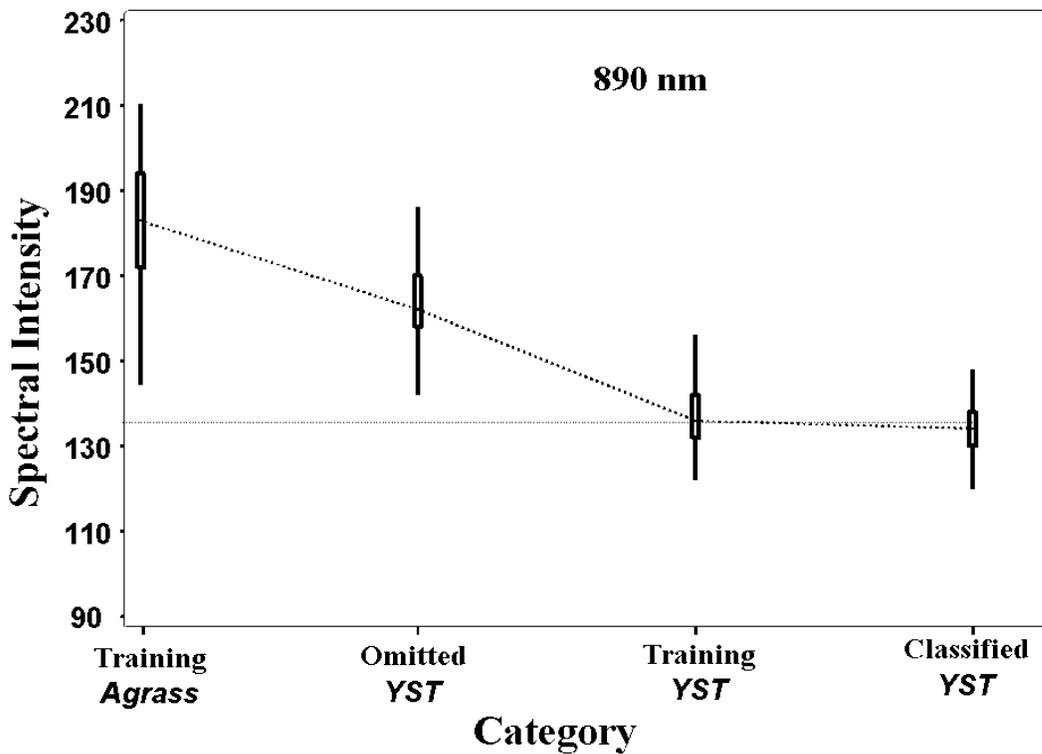

Figure 3. Box plots comparing the spectral response at 890 nm for annual grasses (AGRASS) and yellow starthistle (YST) for various classes.
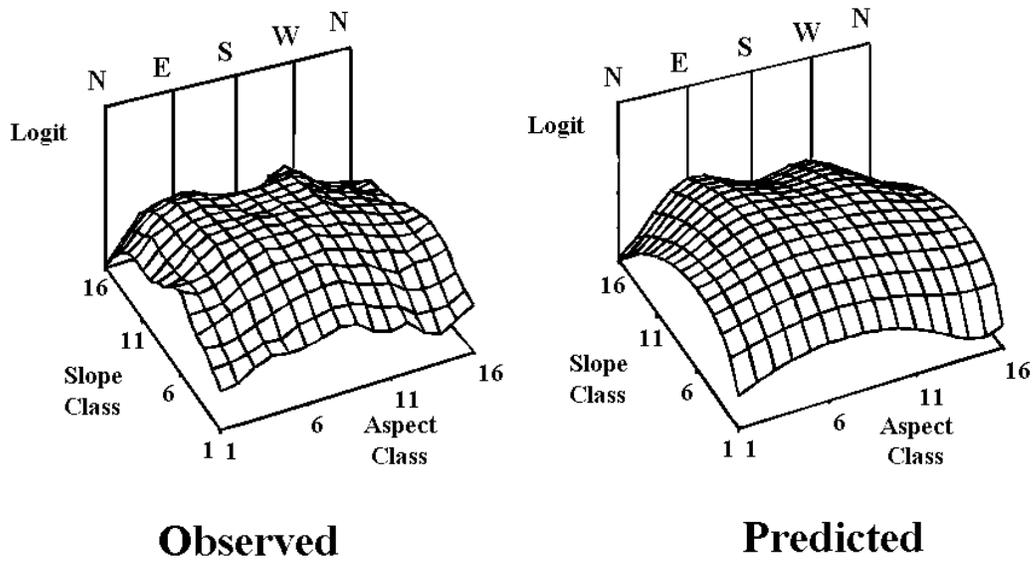
Figure 4. Observed and predicted logit surfaces for the mixed range yellow starthistle landscape model.
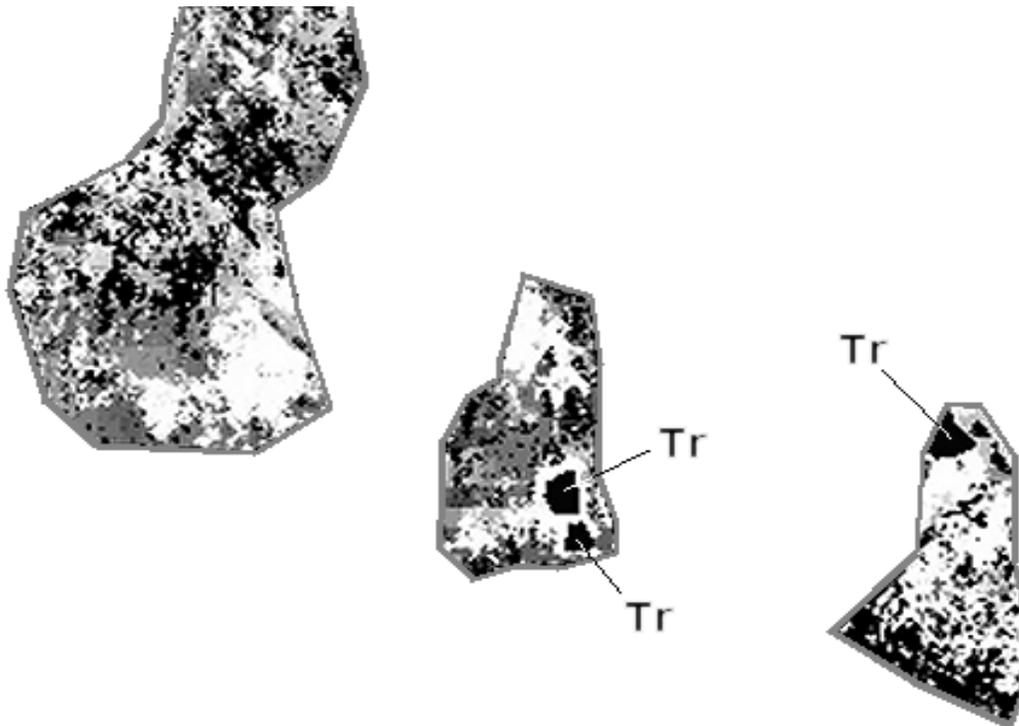

Figure 5. Posterior probability classification for yellow starthistle assuming a spatial prior based on the yellow starthistle landscape prediction model. The probabilities are indicated on a gray scale with black=0.0 and white = 1.0. Areas marked as **Tr** were training sites and were not included in accuracy assessment.
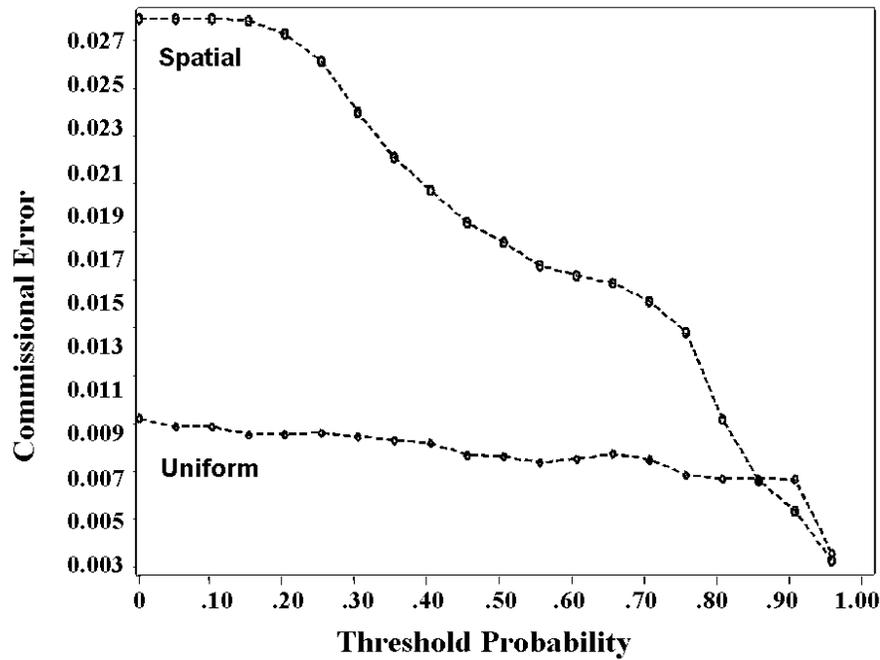
\Figure 6. Effect of changing cut-off threshold probabilities on omissional error rates assuming the uniform prior (dashed line) and the spatial prior (solid line).
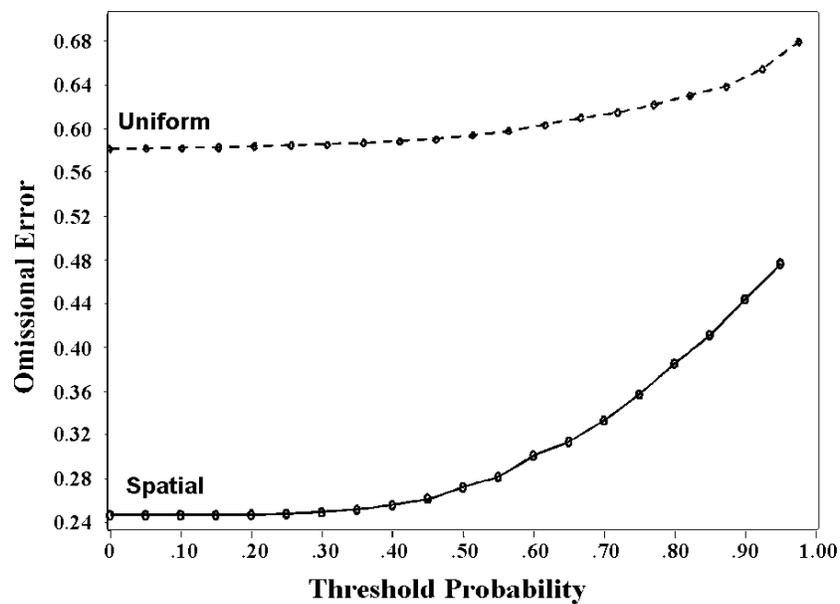

Figure 7. Effect of changing cut-off threshold probabilities on commissional error rates assuming the uniform prior (dashed line) and the spatial prior (solid line).