Conference on Applied Statistics in Agriculture   2006 - 18th Annual Conference Proceedings

MODELING DISPERSAL OF YELLOW STARTHISTLE IN THE CANYON GRASSLANDS OF NORTH CENTRAL IDAHO

Bahman Shafii

William J. Price

Timothy S. Prather

Lawrence W. Lass

Derek Howard

See next page for additional authors

Author Information

Bahman Shafii, William J. Price, Timothy S. Prather, Lawrence W. Lass, and Derek Howard

MODELING DISPERSAL OF YELLOW STARTHISTLE IN THE CANYON GRASSLANDS OF NORTH CENTRAL IDAHO

Bahman Shafii [1,2], William J. Price[1], Timothy S. Prather[2],
Lawrence W. Lass[2], Derek Howard[1]

[1]Statistical Programs
[2]Department of Plant, Soil, and Entomological Sciences
College of Agricultural and Life Sciences
University of Idaho, Moscow, ID

Abstract

Yellow starthistle is an invasive plant species that reduces productivity and plant diversity within the canyon grasslands of Idaho. Early detection of yellow starthistle and predicting its spread have important managerial implications that could greatly reduce the economic/environmental losses due to this weed. The spread of an invasive plant species depends on its ability to reproduce and disperse seed into new areas. Typically, information on the factors that directly affect a plant's ability to reproduce and subsequently disperse seed is not available or difficult to obtain. Alternatively, topographic factors, such as slope and aspect as well as competitive correlates such as vegetation indices related to plant community biomass could be used to model plant survival and seed movement. In this research, several spatial network models incorporating these variables were considered for the prediction of yellow starthistle dispersal. Models will differed in their assessment of plant movement costs, which can be separated into two processes, survival to reproduction and seed dispersal. The candidate models were evaluated based on their predictive ability and biological relevance. Topographical variables, slope and aspect, were found to be significant contributors to yellow starthistle dispersal models, whereas vegetation indices did not improve the prediction process. The optimal model was applied to an area in central Idaho for predicting the dispersal of yellow starthistle in 1987 given a known 1981 infestation.

I. Introduction

Yellow starthistle (*Centaurea solstitialis* L.) can be found in 23 of the 48 contiguous states (Maddox et al. 1985). Although no economic assessments have been conducted for yellow starthistle, millions of dollars in losses may occur from interference with livestock grazing and forage harvesting procedures as well as lower yield and forage quality of rangelands (Callihan et al. 1982; Smathers, R.L. et al. 1985; Roché 1988; Randall, 1994; Callihan, et al. 1995). Numerous reports have also characterized the toxic effect of yellow starthistle on horses (Cordy 1954; Cheeke and Shull 1985; Lass et al. 1999). The best management of yellow starthistle is obtained by identifying the potential weed problems early, controlling them before they reproduce and spread, and monitoring the sites regularly to maintain adequate follow-up control.

The spread of an introduced species depends on establishment and dispersal (Bakker, et al. 1996). Predicting the spread of species is important to management, potentially allowing managers to adjust their activities to either favor spread of desirable species or discourage spread

of invasive plant species. Models that incorporate dispersal have been useful for adjusting management actions for control of invasive plant species (Higgins and Richardson 1998). Predicting movement of a plant species across a specific landscape would allow prescription of techniques to either facilitate or discourage movement.

Movement of a plant species across a specific landscape would require input of site-specific data to accurately represent how a species may move within that landscape. Spatial data residing in a geographic information system (GIS) should serve the process well since data required to model movement is stored in a spatial fashion. The algorithm within GIS software, commonly referred to as the network analysis, may be applied to plant dispersal modeling in this analysis. A network is superimposed on a landscape with raster images supplying the specified landscape attributes (topographical variables, vegetation indices, etc). The establishment and dispersal models estimate the time the invasive plant infestation would take to grow across the network according to the characteristics encountered at that location.

The objective of this research is to develop and evaluate several potential network models that incorporate topographical factors such as slope and aspect for the prediction of yellow starthistle dispersal.

II. Methods

Network description

Network algorithms are commonly implemented in Geographic Information System (GIS) software packages such as IDRISI, a raster based GIS (Clark Labs, 2004) or Arc Info, a vector based GIS (ESRI, 2002). These network routines model the cost of movement along a predefined network , using common mapping data formats. For example, in the simple network shown in Figure 1, movement from point A to point B has a cost of 0.1 units, while the cost of movement from point B to point C is 0.2. That is, movement through the link B → C is twice as costly as movement through the link A → B. Typically, network costs are represented on a unitless, relative basis, and are problem specific in their interpretation. Summation of link costs across a specified path can be useful for determining the furthest movement distance achievable given a fixed cost "budget" (spread analysis) or, alternatively, the path of least resistance between two points on the network (path analysis).

Cost assignment to links on the network can be positive or negative and reflect the resistance of an object through the network. Resistance is represented in the network in one of two ways: active resistance or passive resistance. An active resistance model (dispersal) assumes that cost values greater than 1.0 are forces and values less than 1.0 are frictions. A practical analogy to this model would be that of walking with (force) or against (friction) a wind. Here, the wind is producing an active resistance. The alternative model, passive resistance, assumes that cost values greater than 1.0 result in friction while values less than 1.0 represent forces. The analogy in this case would be one of walking up (friction) or down (force) a hill. In this example, gravity produces a passive form of resistance. In either scenario, cost values equal to 1.0 produces a neutral level of resistance.

An additional component of the network model is a directional effect. If there is independence between the direction of movement and the direction of the applied friction, i.e. the friction is applied equally in all directions, the network is considered to be isotropic. If, on the other hand, the effective friction varies as the movement direction changes, the network is

Conference On Applied Statistics In Agriculture

referred to as anisotropic. In such systems, the friction information is provided with both a magnitude and direction and the effective friction is determined systematically through a directional function. Figure 2 demonstrates the default directional function for the passive cost model in the IDRISI software package. The effective friction (EF) is given as a power function: $EF = (FR)^f$, where FR is the full magnitude of the friction for a specified location and $f = \cos(\quad)^K$. The angle represents the difference between the direction of movement and the direction of the applied friction. K is a parameter which controls the sensitivity to the difference angle, . A value of $K = 0$ produces an isotropic model while larger values lead to a higher degree of directionality in the model. In Figure 2, the effective friction is at 100%, when the movement direction coincides with the friction direction. As the movement direction changes away from the direction of friction, the effective friction diminishes to 0%, when the two directions are perpendicular to each other. As the difference angle increases, the friction becomes a force reaching -100% at 180 degrees. Recalling the hill walking analogy, when one is walking directly up slope, the force is felt at 100%, while walking along the slope produces no resistance or friction. Walking down hill, however, produces a force which is maximized when walking directly down slope. While the IDRISI software package allows modification of the directional function, only the default form given above was used in these simulations.

Application to Yellow Starthistle

For yellow starthistle, the costs of movement over time (dispersal) may be expressed as a function of both the probability of moving between two points, and the probability of surviving to reproduce once the destination point is reached. Although these components could be modeled separately, the yellow starthistle occurrence model described previously by Shafii, et al (2003) combine both movement aspects. This model gives the probability of yellow starthistle occurrence as:

$$Ps_{ij} = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 Y + \beta_4 Y^2 + \beta_5 \ln(X^2 + Y^2) + e_{ij} \qquad (1)$$

Here, Ps_{ij}, or Ps in general, is the probability of yellow starthistle occurrence at the ith, j^{th} levels of topographic slope and aspect, respectively. The values $X = Slope_j * \cos(Aspect_i)$ and $Y = Slope_j * \sin(Aspect_i)$ represent polar coordinate transformations of the slope and aspect data, $_1$ through $_5$ are regression coefficients, and e_{ij} is an error term under usual regression assumptions, $e_{ij} \sim NID (0, \quad^2)$.. The predicted model for a mixed range land use, for example given in Shafii, et al. (2003) is represented in Figure 3. This model generally maximizes the probability of occurrence at moderate slopes in the southwesterly direction.

If C is the cost of yellow starthistle movement between two points, then $C = f(Ps)$ where Ps is the predicted probability of occurrence given by eq (1) at the destination point. A specified form of f(Ps) should logically follow an inverse relationship, that is higher costs of movement would be associated with lower probabilities of occurrence, and vice versa. In this work, the functional forms:

$$f(Ps) \propto (1 - Ps)$$

$$\text{and} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \textbf{(2)}$$
$$f(Ps) \propto (1/Ps)$$

are used. The above functions meet the inverse relationship requirement in both a linear and a nonlinear manner, respectively.

Using slope and aspect information with eqs (1) and (2), a cost or friction map covering the area of interest can be generated. This cost map along with an initial infestation map become the inputs used by the software for simulating yellow starthistle dispersal. In the case of the anisotropic model, a third map indicating friction directions is also used. The output from the network model consists of a predicted image (map). The output map uses a continuous spectrum to represent the extent or distance that an infestation can reach given a fixed "cost budget". Contours of the predicted infestation spread can then be made by tracing an outline along the common spectrum colors or values of the map.

Assessment and Validation

Network model performance was assessed both graphically and numerically. In the graphical assessments, the predicted yellow starthistle spread was visually compared with known infestation maps for similarities and differences in pattern, trend and infestation extent. Numeric quantification of such characteristics is typically expressed through an error matrix and the associated omissional and commissional error rates. These error measurements represent the proportion of image pixels omitted or committed to the final predicted map, respectively (Congalton 1991; Shafii and Price 2001).

Before applying model assessment, however, the model output had to be modified. The assessment methods above require a discrete predicted map to indicate the strict presence or absence of yellow starthistle. This requires a cutoff or threshold probability level to be defined. Hence, the initial model output image was restricted or "hardened" to specific distances or thresholds (Shafii, et al. 2004). Changes in the thresholds, however, produce corresponding changes in the visual assessments as well as the omissional and commissional error rates. By monitoring these changes as the threshold is varied, a hardening level which minimizes either or both error rates and visually fits the known infestation patterns well may be determined. Changes in other model components, such as the direction of the applied friction or the directional sensitivity parameter, can be monitored in a similar manner to identify values which produce a minimal error condition or a good visual fit.

Once a sufficient hardening level, friction direction, and sensitivity were determined, the model was applied to an independent map region. This new region was selected to have a similar environment as the model development area, known yellow starthistle infestations and minimal impact or disturbance from human activities. Validation on the new region was then carried out using the assessment methods above to determine error rates as well as visual comparisons.

III.   Empirical Results and Demonstration

Data

The data used in this study were taken from a University of Idaho aerial survey of the north-central canyons of Idaho. This survey covered Idaho, Latah, Lewis, and Nez Perce

counties of Idaho in the years 1981 and 1987. The survey observations were used to construct a presence/absense map of yellow starthistle infestations at a resolution of 10m x 10m. The average slope value for yellow starthsitle found in the 1981 survey was $31 \pm 8°$ and the average aspect value was $237 \pm 89°$. Additional USGS Digital Elevation Map (DEM) data for slope and aspect information corresponding to the study areas were also used.   For model development, an area adjacent to the Salmon River which covered the Eagle Creek drainage was selected. Model validation was carried out on an area west of Kamiah, Idaho along Laywer Creek. Both sites were chosen to have known yellow starthistle infestations and minimal ifluences from roads, trails or agricultural disturbances.

Normalized Difference Vegetation Index (NDVI) and Transformed Soil Adjusted Vegetation Index (TSAVI) were calculated from red and near infrared bands of LandSAT 5 data acquired June 30, 1987. Spatial resolution of the LandSAT data was reduced to 10m by 10m during the rectification process. The NDVI and TSAVI values were scaled to match training sites.   NDVI estimates the percent green vegetation cover or healthy vegetation while the TSAVI indicates the amount of living vegetation present (biomass) adjusted for soil background and hence, both indices provide information on competitive ability.

All map data were incorporated into the IDRISI Kilimanjaro GIS package. Figure 4 shows an IDRISI image of the Eagle Creek drainage used for model development. The red areas of the image indicate the known 1981 yellow starthistle infestations. These points were used as the initial starting infestations in the network model. Lighter yellow regions represent the known expansion of yellow starthistle through 1987. These boundaries were used in the model assessment process.

Incorporating the USGS DEM Eagle Creek data (slope and aspect) with eq (1) produced the predicted yellow starthistle occurrence map (Ps) shown in Figure 5. Here, darker black to blue colors indicate lower probabilities of occurrence and brighter yellow and red colors correspond to higher probabilities. The occurrence map was subsequently converted to corresponding cost maps using the forms in eq (2) (Figure 6). Interprettion of the cost map colors is opposite that of the corresponding occurrence maps. Each of the cost maps, i.e. linear and nonlinear, were evaluated separately in the modeling process.

Isotropic Model

Figure 7 demonstrates the output resulting from the isotropic cost model overlayed with the known 1981 and 1987 Eagle Creek yellow starthistle infestations. In this example the model also assumed the cost function, $C = (1/Ps)$. The predicted distances (colors) show a general trend of darker values (lower costs of movement) close to the initial 1981 yellow starthistle infestations which change to yellow and red colors (higher movement costs) as the distance increases away from the 1981 boundaries. An example of model assessment is shown in Figure 8 where hardening levels were set from 100 to 700 units. Generally, the commisional error rate, i.e. the number of pixels eroneously committed to yellow starthistle, increases as the hardening level increases and the predicted infestation boundary expands. In contrast, the omissional error rate, or number of pixels omitted from yellow starthistle classification, decreases as the hardening level increases. Appropriate or acceptable levels of each error rate will depend on both the intended use and user of the predicted map. While this determination is subjective, the point at

which both error rates are minimized was selected in this study as a basis for model comparisions. In the case of the isotropic model above, a hardening level of 360 units produces a combined minimal error rate of 27%. The initial prediction (dark red) along with the initial 1981 infestation (light red) and the 1987 infestation (yellow) are shown in Figure 9. Following a similar modeling and assessment process for the cost function C = (1 - Ps), the prediction in Figure 10 was obtained. While both models produce similar error rates (27 - 28%), the hardening level used for the (1 - Ps) model was 1500 units. Although the hardening levels do not have a direct interpretation, they do represent a measurement of model efficiency. In this case, the later model C = (1 - Ps) must be extrapolated further to achieve equivalent results to that of the first model, i.e. 1500 vs 360 units. Thus, the cost function C = 1/Ps was selected as the better model and will be used in the subsequent analyses. Nonetheless, neither of the isotropic models covered the known 1987 infestation well, suggesting an anisotropic solution may be more appropriate.

Anisotropic Model

For the anisotropic model, an image which gives directional friction information is required. Initially, the topographic map of aspect was used for this purpose. That is, the model assumed friction direction to coincide with aspect. The directional sensitivity for this model was set to K = 1, the default value in the IDRISI software package.

The output from the anisotropic model is given in Figure 11. Under the anisotropic conditions, the cocurrent error rate was reduced substantially to 11.8% given a hardening level of 60 units. The general pattern of predicted infestation also coincides closely to the observed 1987 infestation boundaries. Overall, the anisotropic model performs much better than the corresponding isotropic model. The incorporation of directional information improved the prediction by reducing both the error rates and hardening levels. Furthermore, the directional input was topographic aspect, suggesting that aspect may be an important factor in yellow starthistle movement.

Model Enhancement

While directional information greatly improved the model predictions, the relationship of aspect and movement was explored further. Specifically, the initial directional input assumed friction direction to coincide with aspect. In order to investigate the effects of friction direction on model performance, the angle of the friction was varied in 10 degree increments in a clock-wise fashion from 0 degrees (aligned with aspect) to 350 degrees. Figure 12 shows the effect of this rotation on the error rates. Unlike the changes in the hardening level, variations in the friction direction produced three points where the omissional and commissional error rates were equal. Overall, however, one angle of rotation, 50 degrees, resulted in a minimal value of 11.8% for both error types. Given this rotation, the hardening level was also reduced to 50 units. The 50 degree frictional direction may be of biological relevance. Observational evidence from the yellow starthistle survey shows that the average aspect for yellow starthistle occurrence is 237 degrees, which is approximately 180 degrees from the direction indicated by the model.

Another model component examined was the directional sensitivity parameter, K. Figure 13 shows the effects of changing K from 0 to 10, given the previous model. While the

commissional error rate appears to stabilize at higher values of K, the ommissional error rate does not. A minimal error rate of 11.8% is achieved, however, at K=2. In addition, this value of K reduces the hardening level further to 48.5 units.

Further enhancement of the model was considered through the addition of the aforementioned vegetation indices, NDVI and TSAVI, as competitive correlates related to plant community biomass. These measures were incorporated into the modeling process both as a direct adjustment to the predicted probability of yellow starthistle occurrence as well as an indirect adjustment to the directional frictions. However, their inclusion had little or no influence on the resulting prediction models. This may be due to the known correlation between aspect and the annual grass community where yellow starthistle thrives. Since the predicitive probablity model (1) is based on observed yellow starthistle infestations in these plant communities, it is likely that the specified dispersal model may already be accounting, as a proxy, for the effect of these factors.

Figure 14 shows the predicted results of the final model. Like the previous model, the predicted boundaries follow the 1987 infestation pattern well, with some additional improvements.

Validation

The area used for validation of the model, Laywer Creek, was located several kilometers east of the Eagle Creek drainage, near Kamiah, Idaho. Although Laywer Creek is surrounded by agricultural fields and has several roads and trails present, the area used for validation was isolated from these features. The creek had several identifiable yellow starthistle infestations in1981 to be used as initial starting points for the simulation. Figure 15 shows the predicted and known infestations for this region. While the model performed reasonable in most cases, there were some areas of over prediction (indicated by yellow) leading to an omissional error rate of 21.7%. The commissional error rate was relatively lower at 5.7%. In general, the model did well in predicting the spread of yellow starthistle in the validation area.

IV. Concluding Remarks & Future Work

Prediction of yellow starthistle dispersal has important economic and managerial advantages. It would allow strategic planning efforts to focus on areas at highest risk, and identification of defensible boundaries for optimizing limited management resources. The network model provided a reasonable assessment of dispersal of yellow starthistle given the known infestation in the specified area, and validated well in an independently larger area. Potential future models should be flexible to investigate the effects of topographical factors (e.g. up slope versus down slope), climatic effects (e.g. wind, precipitation), physical inputs (e.g. insolation, soil properties), human disturbances (e.g. roads, trails, buildings), or other biological/physical metrics. These additional components may be incorporated into the modeling process directly through adjustments to the predicted probabilities or indirectly through modification of the friction direction. It would also be useful to develop alternative means for assessing the prediction errors using parametric or nonparametric simulations.

V. References

[ESRI] software, 2002. version 8.2. four parts: ArcINFO work station, ArcMap, ArcCatalog, and ArcToolbox. NY:Environment Systems Research Institute.

Bakker, J. P., P. Poschlod, R. J. Strykstra, R. M. Bekker, and K. Thompson. 1996. Seed banks and seed dispersal: important topics in restoration ecology. Acta Botanica Neerlandica 45:461-490.

Callihan, R.H., R.L Sheley, and D.C Thill. 1982. Yellow starthistle: identification and control. Current Info. Series No. 634, 4 pp. University of Idaho, College of Agriculture.

Callihan, R. H., L. Smith, J. P. McCaffrey, and E. Michalson. 1995. Yellow starthistle management for small acreages. Moscow, ID: University of Idaho Cooperative Extension System Bulletin CIS 1025, 8p.

Checke, P.R. and L.R. Shull. 1985. Other plant toxins and poisonous plants. Ch. 11. Pages 358-392. In, Natural Toxicants in Feeds and Poisonous Plants. The Avi Publ. Co., Westport, CT.

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

Cordy, D.R. 1954. Nigropallidal encephalomalacia (chewing disease) in horses on rations high in yellow star thistle. Proc., Am. Vet. Med. Assoc. 91:149-154.

Higgins, S. I. and D. M. Richardson. 1999. Predicting plant migration rates in a changing world: the role of long-distance dispersal. The American Naturalist 153:464-475.

IDRISI GIS Software, 2004, Clark Labs, Clark University, 950 Main Street, Worcester, MA 01610-1477

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.

Maddox, D.M., A. Mayfield, and N.H. Poritz. 1985. Distribution of yellow starthistle (Centaurea solstitialis) and Russian knapweed (Centaurea repens). Weed Science 33(3):315-327.

Roché, C.T. and B.F. Roché, Jr. 1988. Distribution and amount of four knapweed (Centaurea L.) species in eastern Washington. Northwest Science 62:242-253.

Randall, J.M. 1994. Weeds and the preservation of natural areas. Proc.,California Weed Science Society 46:143-154.

Shafii, B. and W. J. Price. 2001. Comparing binomial, bootstrap and Bayesian estimation methods in assessing the agreement between classified images and ground truth data. Pages 70-84 in Proceedings of the thirteenth conference on Applied Statistics in Agriculture. Manhattan, KS: Kansas State University.

Shafii, B., W. J. Price, T. S. Prather, L. W. Lass, and D.C. Thill. 2003. Predicting the likelihood of yellow starthistle occurrence using landscape characteristics. Weed Science, 51:784-751.

Shafii, B., J. W. Price, T. S. Prather, L. W. Lass, and D. C. Thill. 2004. Using landscape characteristics as prior information for Bayesian classification of yellow starthistle. Weed Science. 52:948-953.
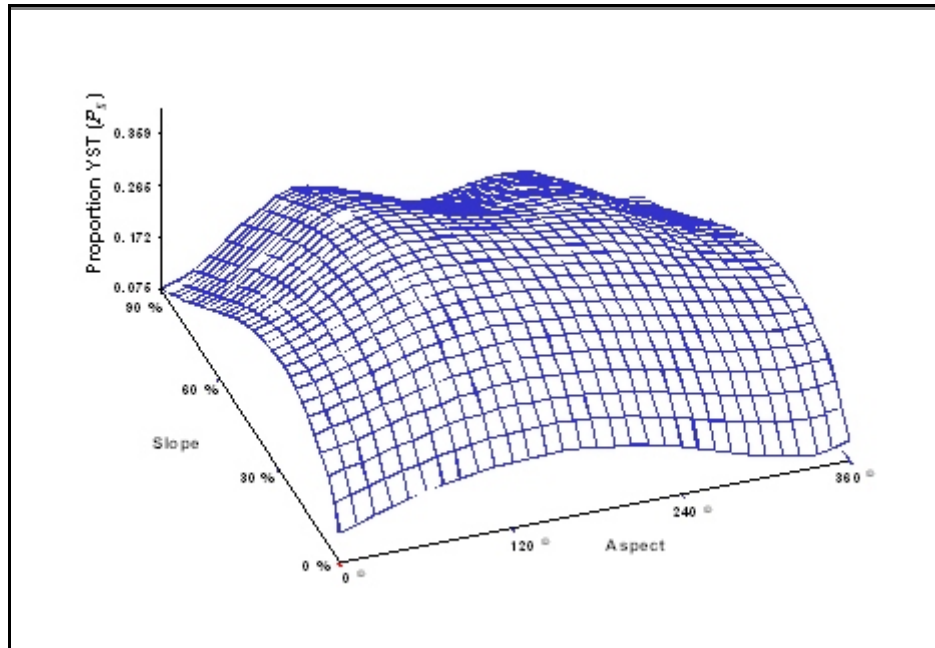
Smathers, R.L., R.O. Brooks, and E.L. Michalson. 1985. Controlling yellow starthistle economic considerations. Bull. Idaho Agric. Exp. Stn. 8(650):1-8.
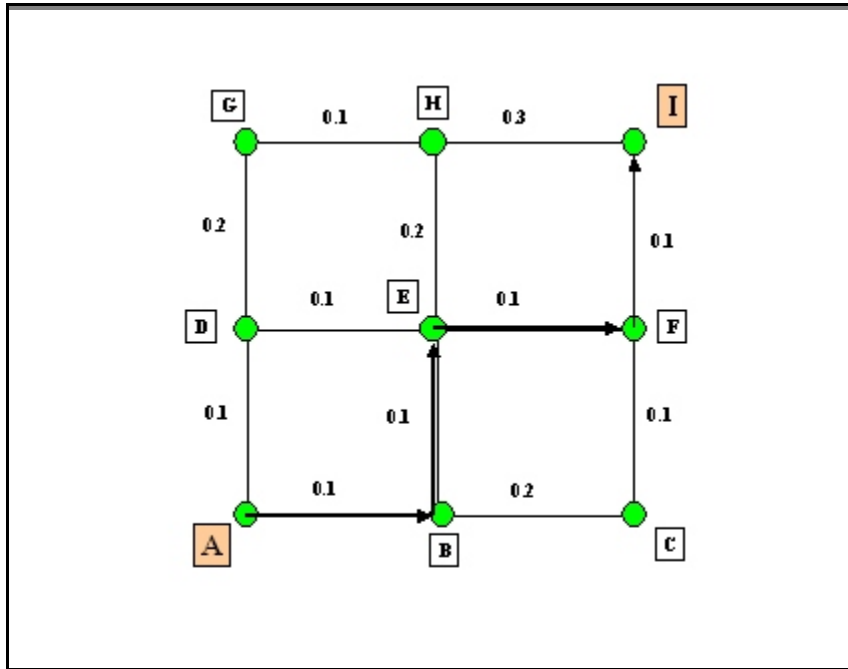
Figure 1. Simple 9 node network diagram with associated links
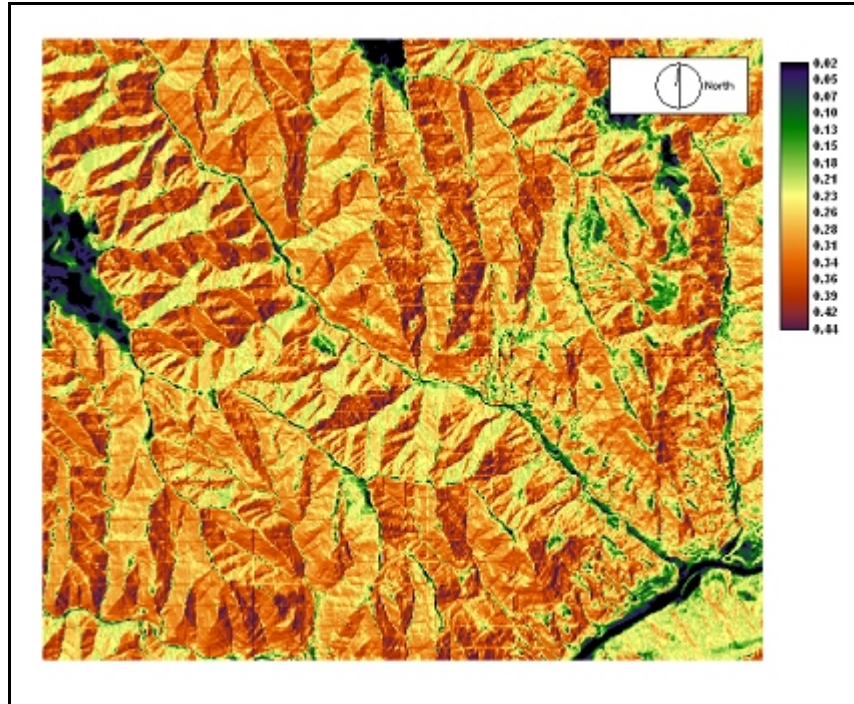and costs.
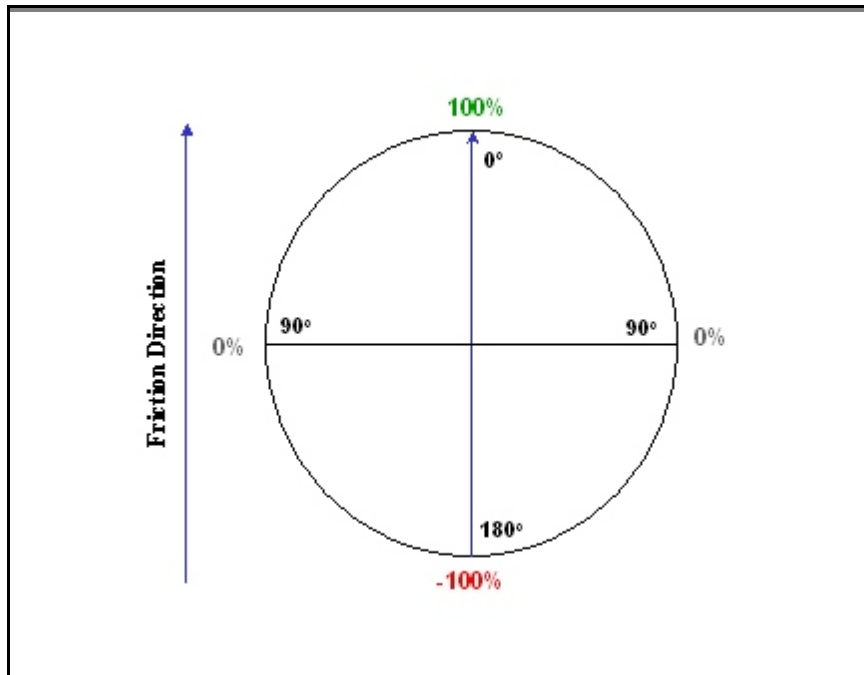
**Figure 2**. Effective friction relative to movement direction.

Figure 3. Example of the predicted aspect-slope response surface for the mixed range land use.


Figure 4. IDRISI image of the Eagle Creek drainage. The 1981 infestation used for model development is shown in red while the 1987 infestation used for model assessment is shown in yellow.

Figure 5. Predicted yellow starthistle occurrence map. Darker
black to blue colors represent lower likelihood of occurrence
while yellow to red colors represent higher likelihoods.
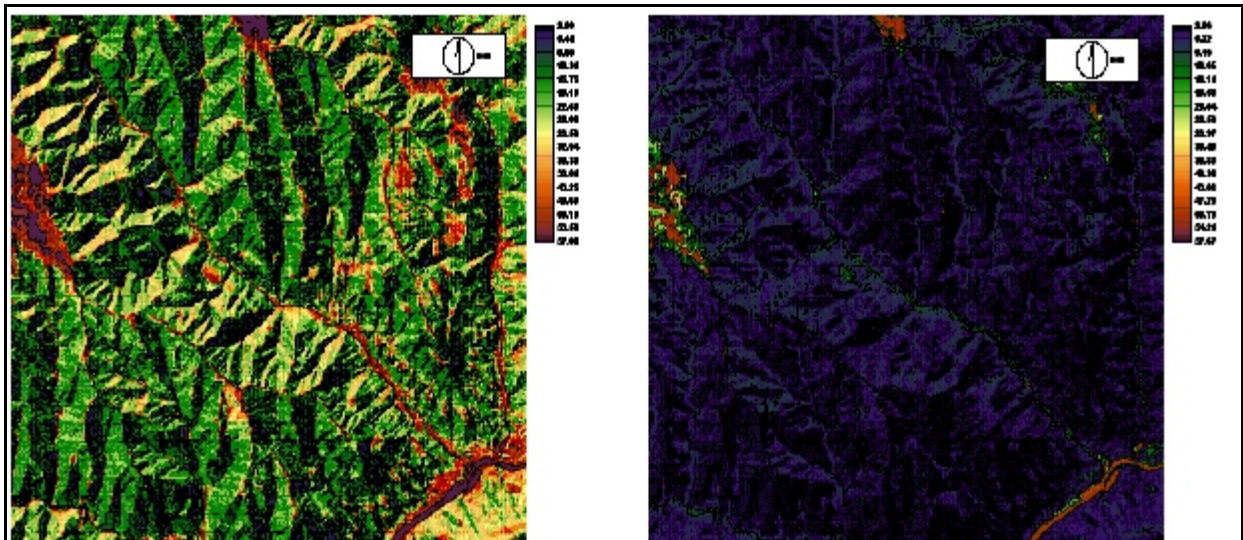
Figure 6. Transformed cost maps for the linear (left) and the nonlinear (right) cost functions.
Darker black to blue colors represent higher costs while yellow to red colors represent lower
costs.
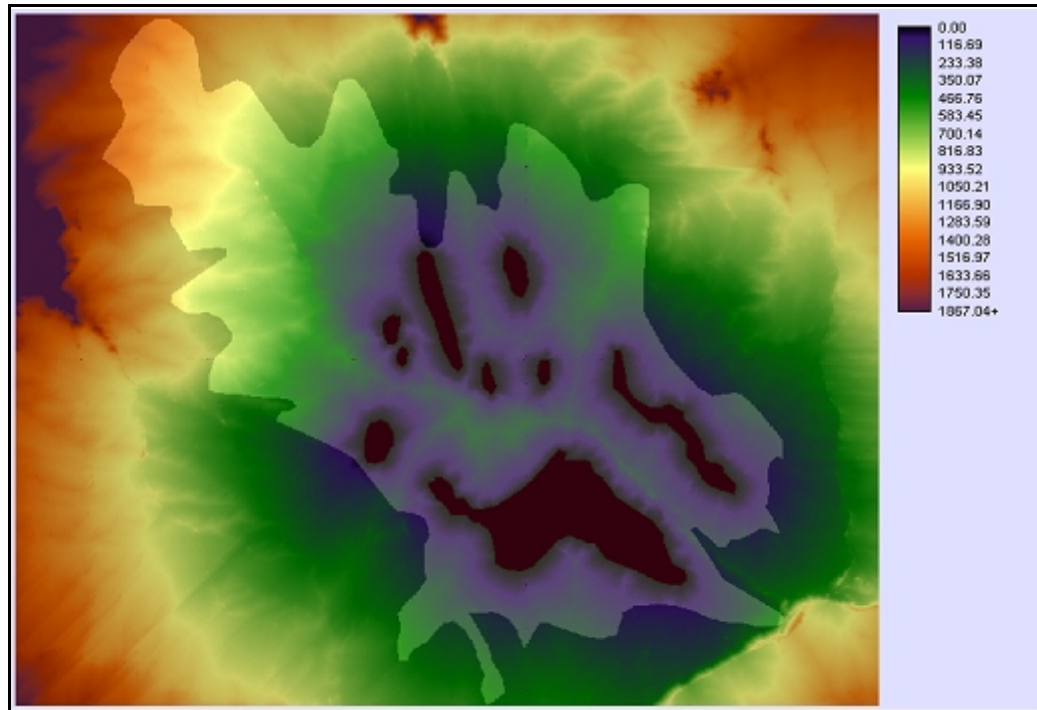
Figure 7. Predicted map of the isotropic cost model (blue to orange spectrum) overlaid with the the known initial 1981 (dark red) and subsequent 1987 (lighter yellow) starthistle infestations.
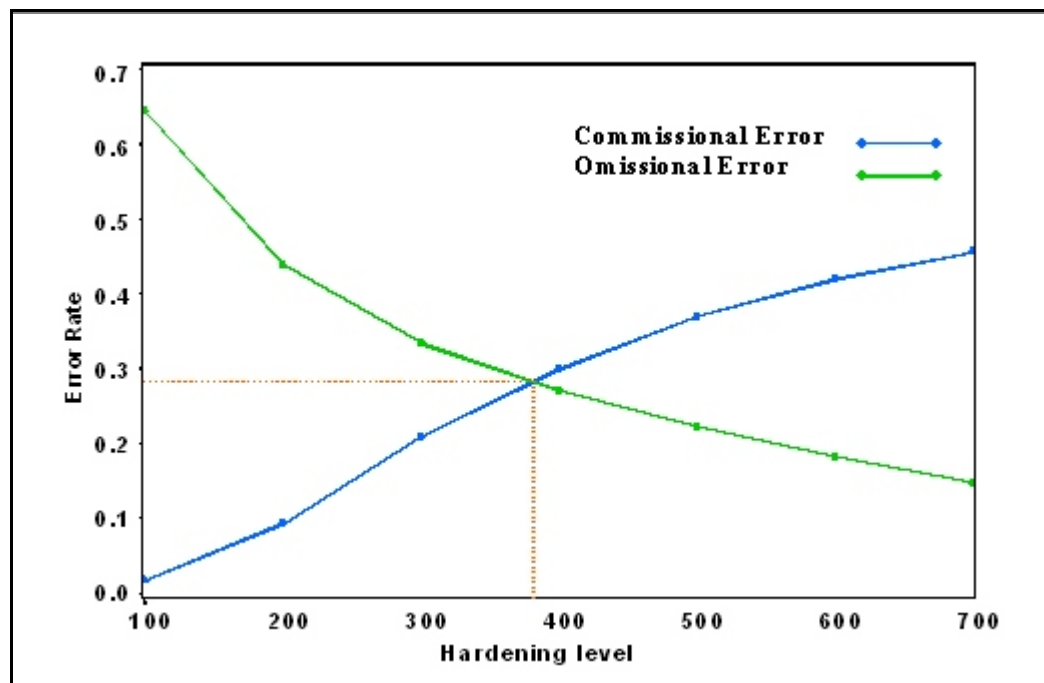
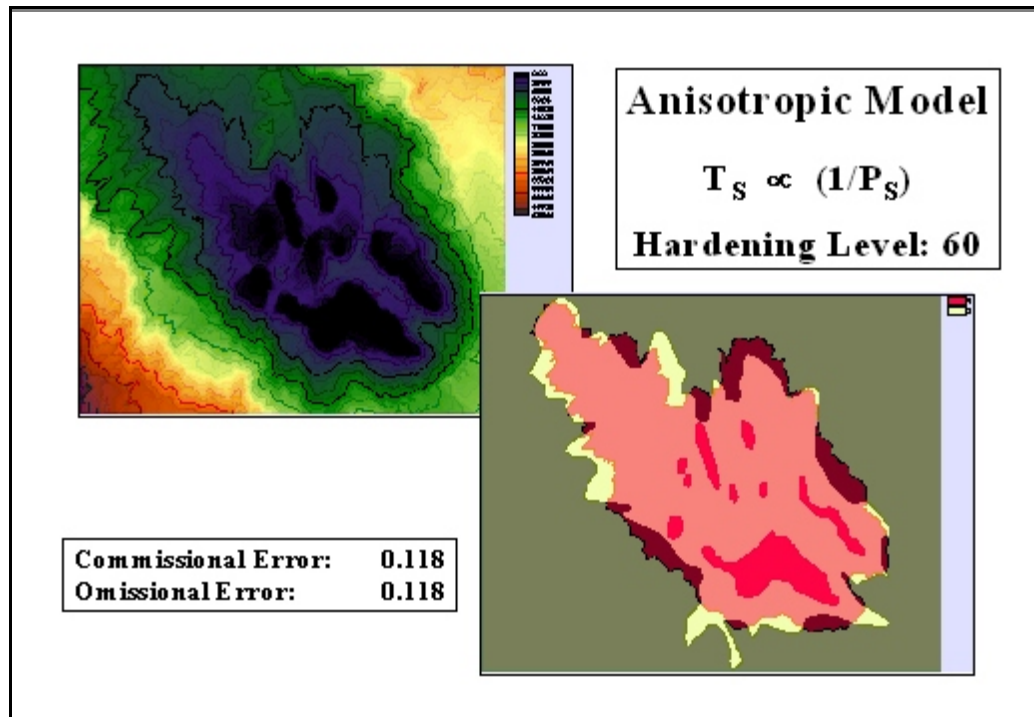**Figure 8**. Omissional and commisional error rates for various hardening levels of the isotropic cost model.
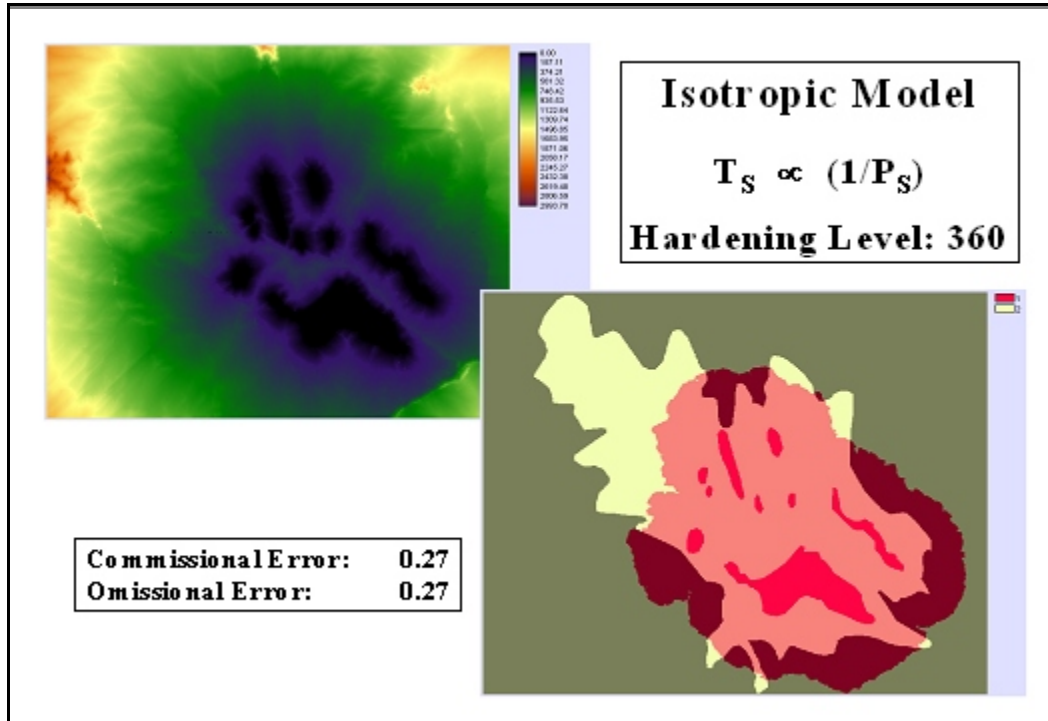
Figure 9. The isotropic cost model prediction assuming the nonlinear cost function (pink and maroon colors). The1981 initial infestation (red) and the later 1987 assessment infestation (yellow) are overlaid for comparison.
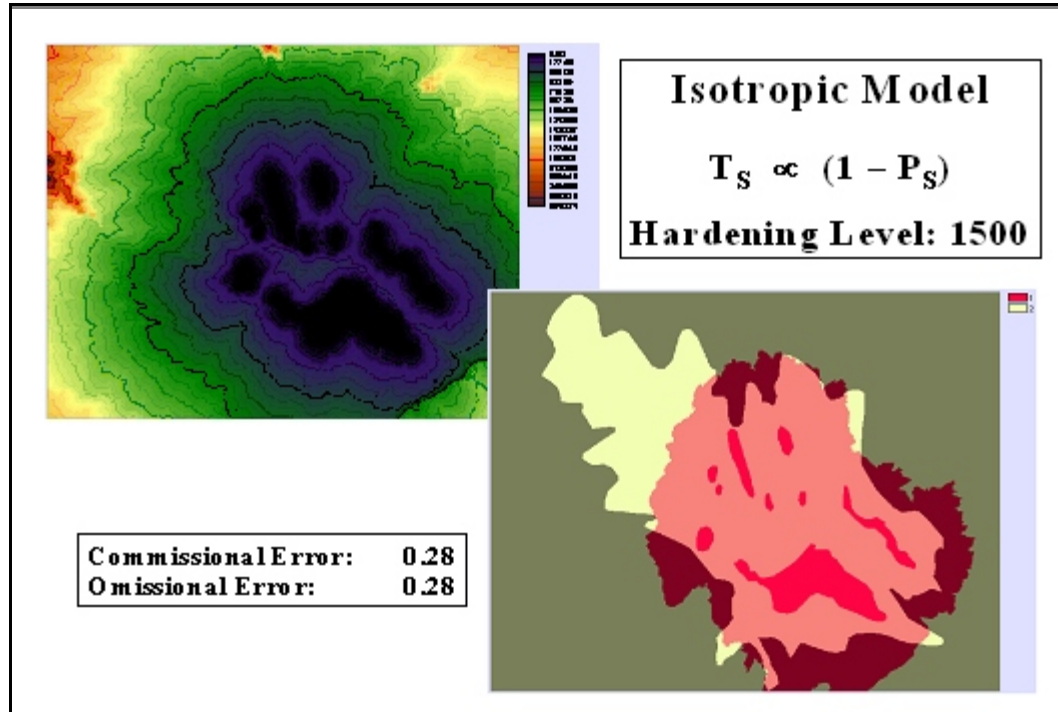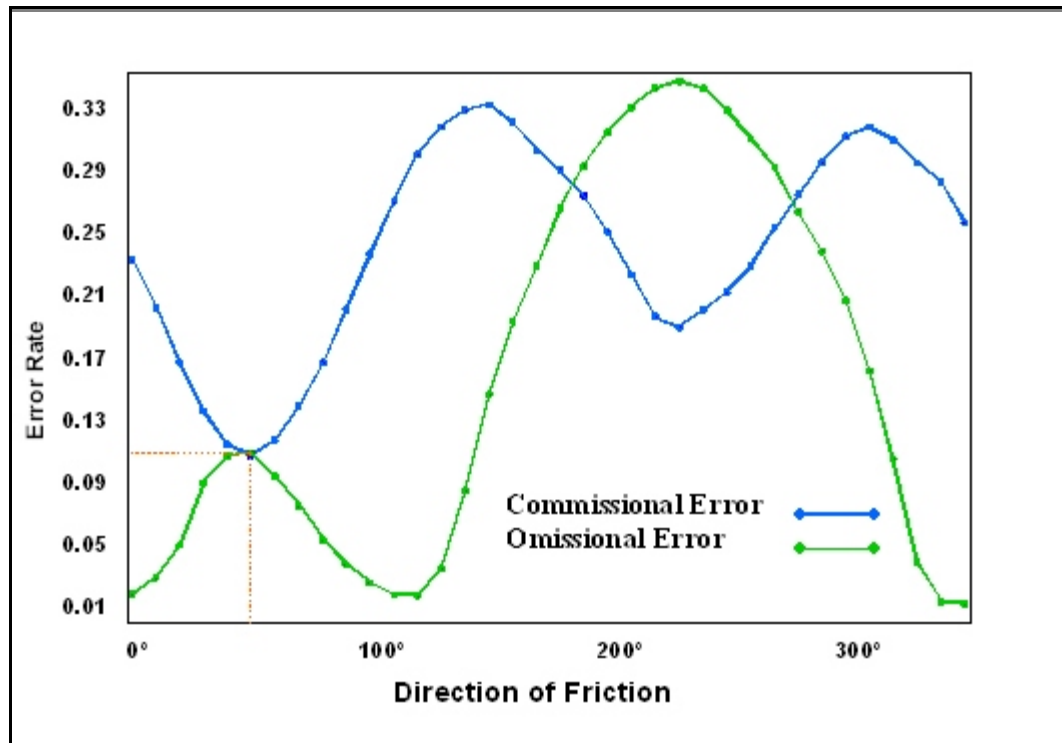

Figure 10. The isotropic cost model prediction (pink and maroon colors) assuming the linear cost function. The1981 initial infestation (red) and the later 1987 assessment infestation (yellow) are overlaid for comparison.
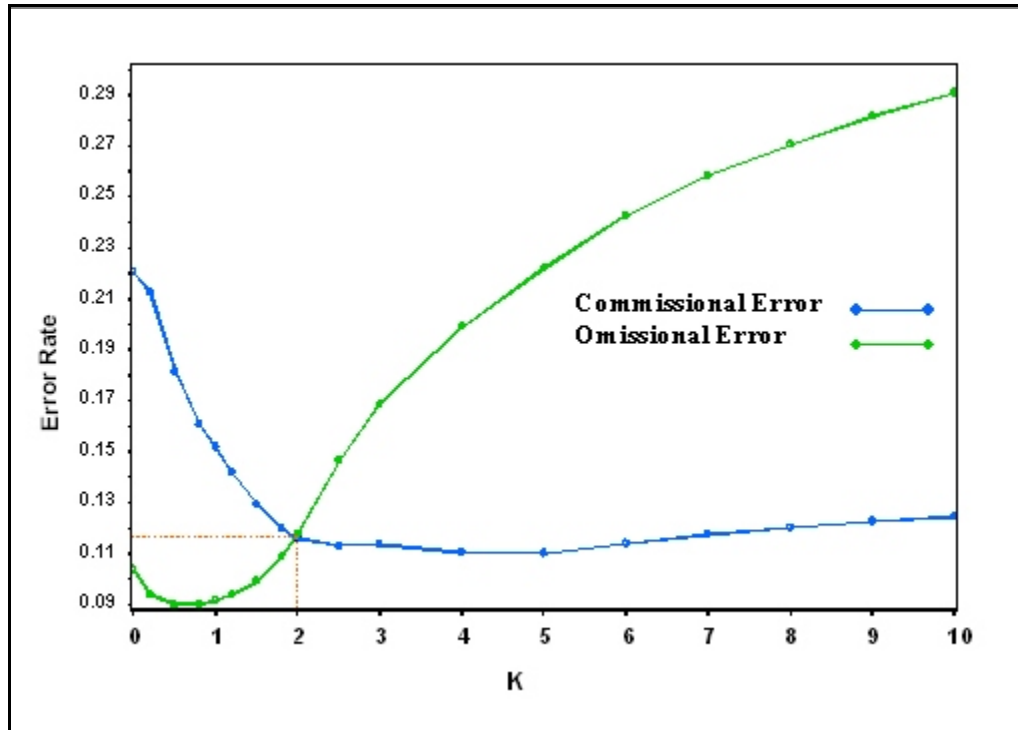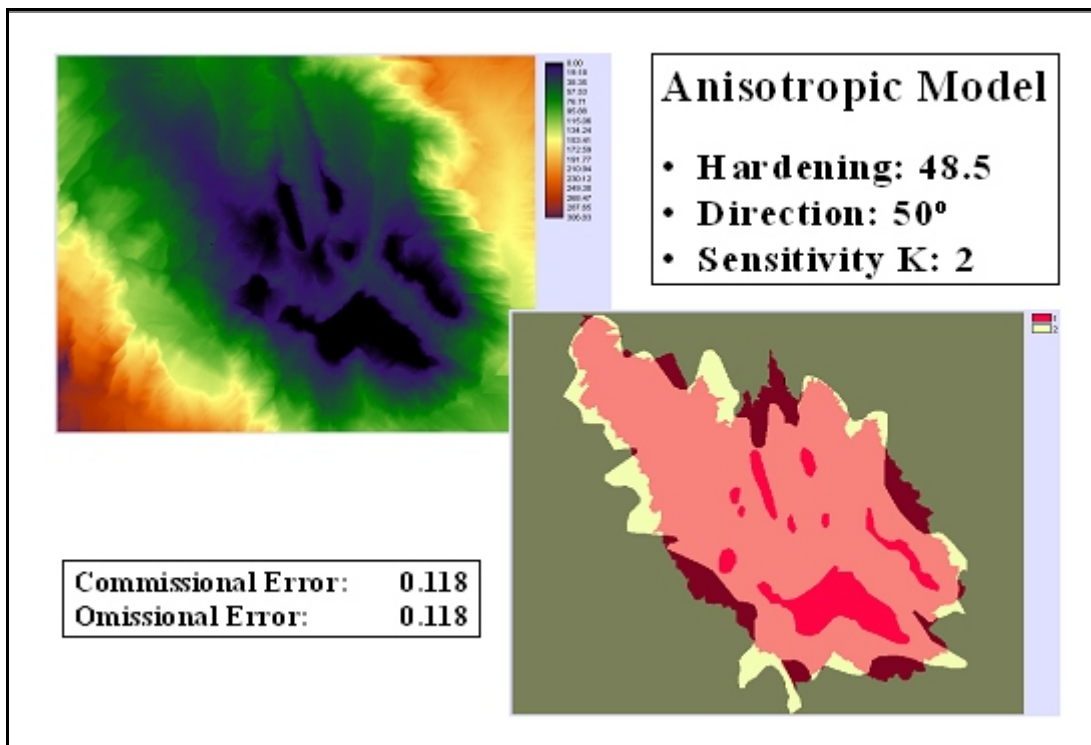
Figure 11. The anisotropic cost model prediction (pink and maroon colors) assuming the nonlinear cost function. The1981 initial infestation (red) and the later 1987 assessment infestation (yellow) are overlaid for comparison.


Figure 12. Omissional and commissional error rates resulting from a rotation of the friction direction.

Figure 13. The effect of the sensitivity parameter, K, on the omissional and commissional error rates.

Figure 14. The final anisotropic cost model prediction (pik and maroon colors) assuming the nonlinear cost function. The1981 initial infestation (red) and the later 1987 assessment infestation (yellow) are overlaid for comparison.
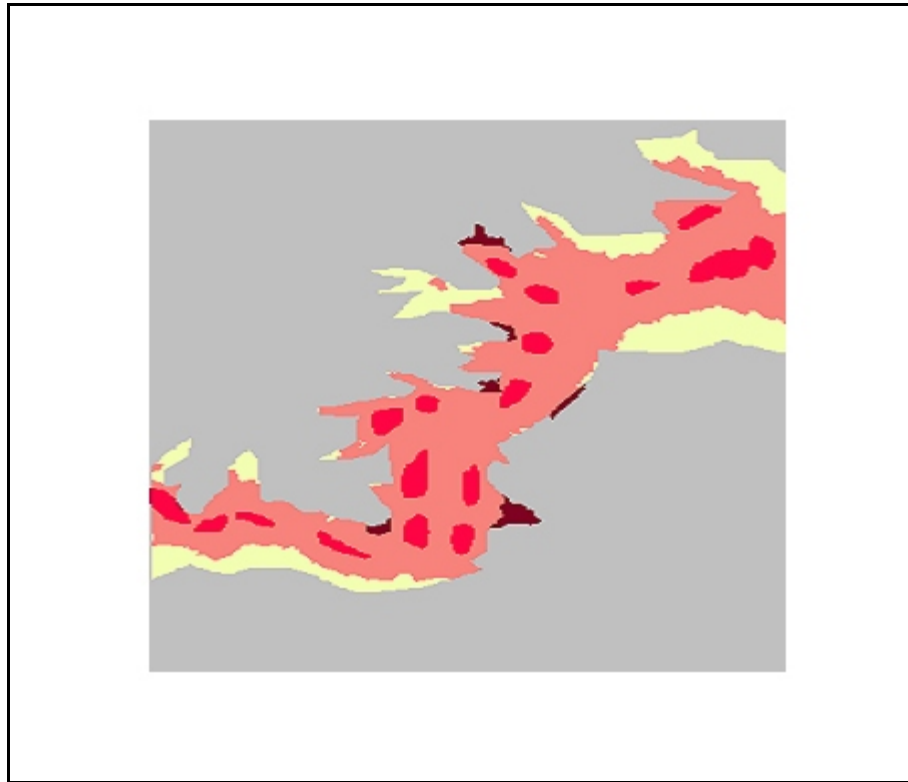
Figure 15. The anisotropic cost model validation (pink and maroon colors).The1981 initial infestation (red) and the later 1987 assessment infestation (yellow) are overlaid for comparison.