Cohen's Kappa and Classification Table Metrics 2.0

An ArcView 3.x Extension Developed To Verify Spatially Explicit Predictive Models Through A Multi-Criteria Selection Process

Jeff Jenness[1] and J. Judson Wynne[2]

Introduction

Spatially explicit models have various applications in resource management, including the development of vegetation and wildlife-habitat predictive surface maps. Appropriate applications of these models are impossible without informed approaches to model development and accuracy assessment of resultant data products. In the absence of incisive model development and error analysis, spatially explicit models may be applied in ways which confound, rather than illuminate, our understanding of vegetation land cover and wildlife habitat.

Accuracy assessment provides a means of gauging model performance, and thus may serve to elucidate our understanding of our developed predictive models. End users who conduct accuracy assessment are also provided with important information regarding model reliability and suitability of the modeling process (Csuti and Crist 1998; Drost et al. 1999).

The Kappa Analysis extension will provide end users with a packaged approach for accuracy assessment, including the Kappa statistic as well as several additional metrics, to be used for gauging model performance. When multiple competing models are available, these metrics can be used to quantitatively compare and identify the "best" model within a multi-criteria model selection process.

Kappa Analysis strives to raise the bar for accuracy assessment and provide a quantative approach to making model comparisons. We hope users of this product agree.

Kappa (or KHAT) $\hat{K} = \dfrac{n\sum_{i=1}^{k} n_{ii} - \sum_{i=1}^{k} n_{i.}n_{.i}}{n^2 - \sum_{i=1}^{k} n_{i.}n_{.i}}$

with Variance $\hat{\mathrm{var}}(\hat{K}) = \dfrac{1}{n}\left(\dfrac{\theta_1(1-\theta_1)}{(1-\theta_2)^2} + \dfrac{2(1-\theta_1)(2\theta_1\theta_2-\theta_3)}{(1-\theta_2)^3} + \dfrac{(1-\theta_1)^2(\theta_4-4\theta_2^2)}{(1-\theta_2)^4}\right)$

where:

$$\theta_1 = \frac{1}{n}\sum_{i=1}^{k} n_{ii}$$

$$\theta_2 = \frac{1}{n^2}\sum_{i=1}^{k} n_{i.}n_{.i}$$

$$\theta_3 = \frac{1}{n^2}\sum_{i=1}^{k} n_{ii}(n_{i.}+n_{.i})$$

$$\theta_4 = \frac{1}{n^3}\sum_{i=1}^{k}\sum_{j=1}^{k} n_{ij}(n_{j.}+n_{.j})$$

j = Columns

Reference Data: Actual or Field-checked

i = Rows — *Classification Data: Predicted from Model*

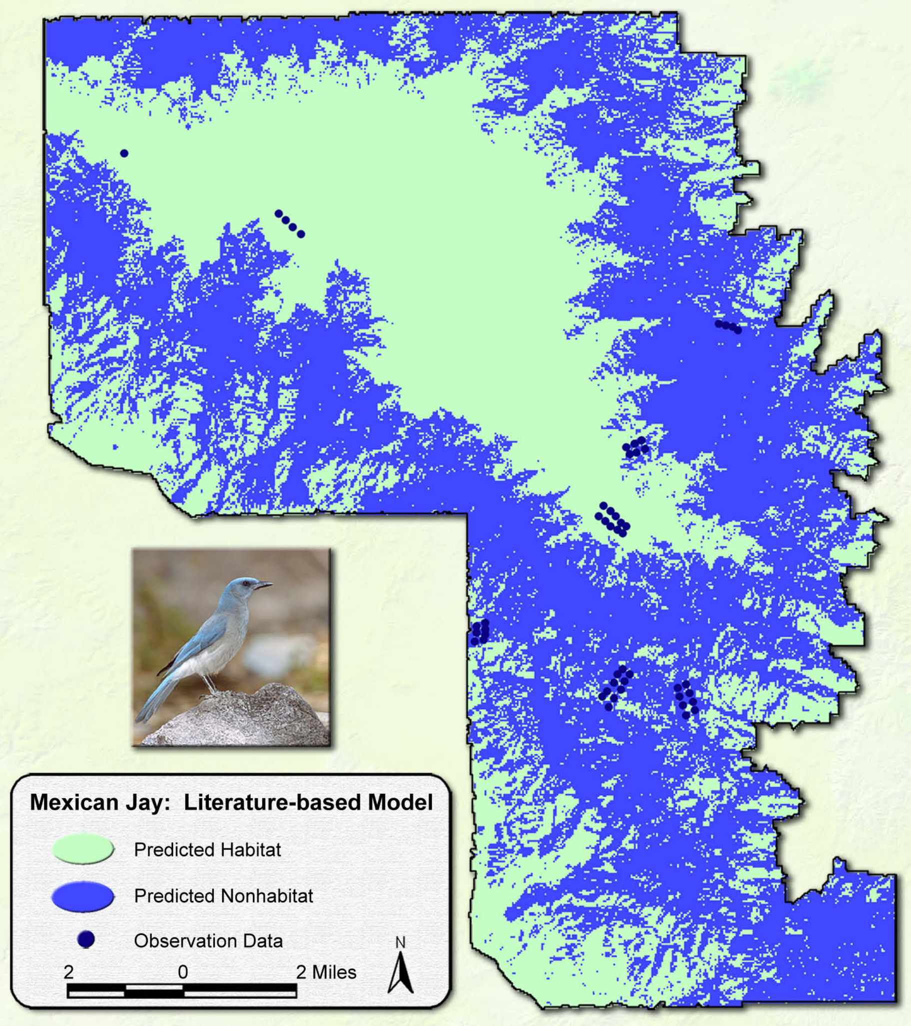
	j_1	j_2	j_k	$n_{i.}$ *(Row Totals)*
i_1	n_{11}	n_{12}	n_{1k}	$n_{1.}$
i_2	n_{21}	n_{22}	n_{2k}	$n_{2.}$
i_k	n_{k1}	n_{k2}	n_{kk}	$n_{k.}$
$n_{.j}$ *(Column Totals)*	$n_{.1}$	$n_{.2}$	$n_{.k}$	$n_{..} = n$

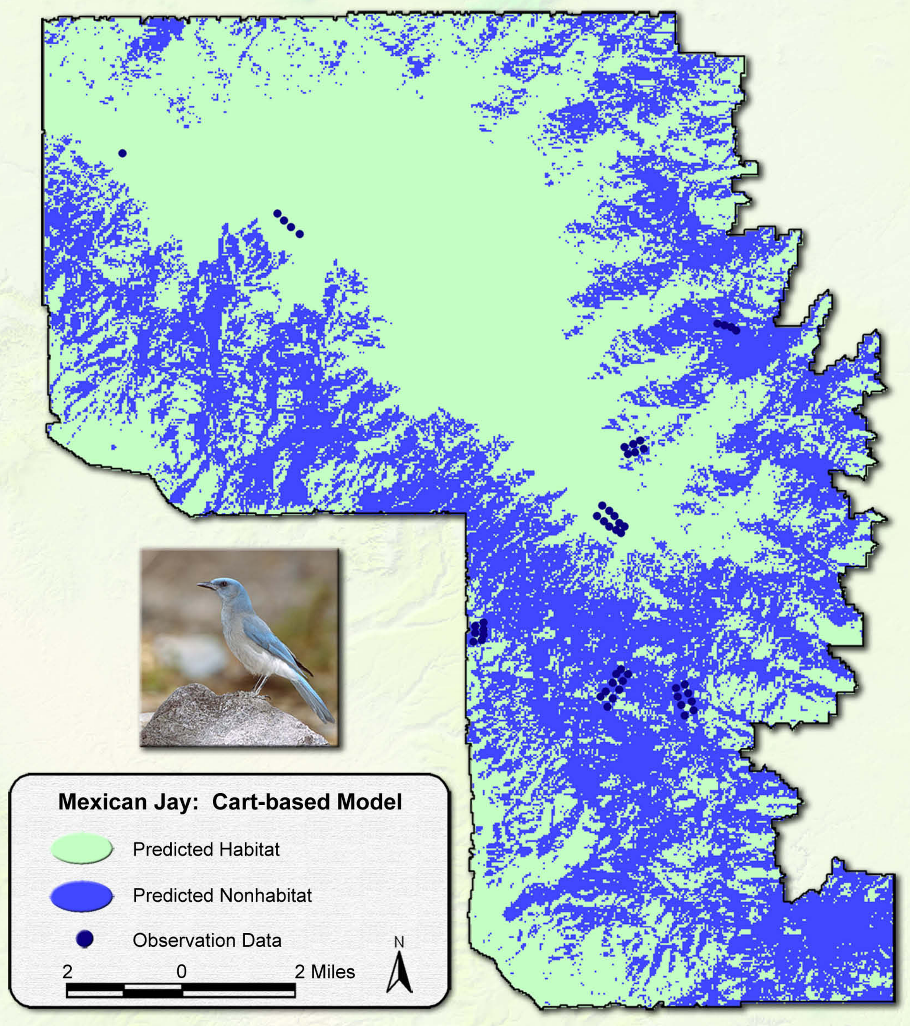
(Adapted from Congalton and Green 1999, p. 47)

Comparing Models

Metric	Literature-based	Classification-tree based
Overall Accuracy	56.8	**63.6**
Misclassification Rate	43.18	**36.36**
K_{HAT}	0.136	0.22
P-value	0.229	0.09
Sensitivity (Absence)	0.545	**0.636**
Sensitivity (Presence)	0.636	0.636
Specificity (Absence)	0.636	0.636
Specificity (Presence)	0.545	**0.636**
Positive Predictive Power (Absent)	0.818	**0.84**
Positive Predictive Power (Presence)	0.318	**0.368**
Commission (Absence)	0.364	0.364
Commission (Presence)	0.455	**0.364**
Omission (Absence)	0.455	**0.364**
Omission (Presence)	0.364	0.364

Bold text indicates better performance by model.





This Kappa Analysis ArcView extension provides several metrics that describe predictive models, therefore providing a useful method for identifying the model with the highest potential performance even when there are no statistically significant differences between models.

For example, we generated and compared a classification tree based model and a literature-based model for Mexican jay (*Aphelocoma ultramarina*) habitat in the Pinaleños Mountains, Arizona. The classification tree based model was developed using a 1993-95 retrospective dataset collected by Dr. William M. Block, USDA Forest Service, Rocky Mountain Research Station, Flagstaff. Literature-based information was derived from the Birds of North America species account and other peer-reviewed sources.

We identified the model with the best performance using multiple metrics from the error matrix, including highest overall accuracy, lowest misclassification rate, highest K_{HAT} statistic, *P*-value (significant at ≤ 0.05), highest sensitivity and specificity, highest positive predictive power, and lowest commission and omission error rates (Fielding and Bell 1997, Congalton and Green 1999).

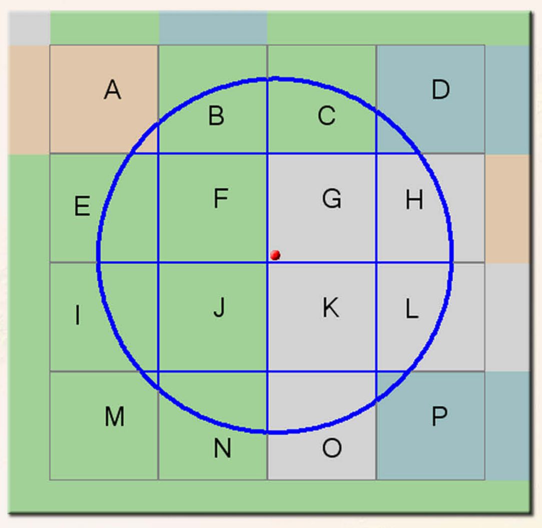
A statistical comparison of the models suggests no significant difference exists between the literature-based and classification tree based models (Z=0.335, p = 0.369). Furthermore, neither model was statistically significant. However, using a multi-criteria model selection process, the classification tree-based model performed better than the literature-based model (see Table above). The classification tree-based model had the highest overall accuracy (63.6%), lowest misclassification rate (0.364), and highest Kappa value (K_{HAT} = 0.22). Additionally, this model had highest specificity for predicting presence (0.636), highest model sensitivity for absence (0.636), highest predictive power for absence (0.84) and presence (0.368), lowest commission rates for presence (0.364) and lowest omission rates for absence (0.364).

Incorporating Locational Uncertainty

Locational uncertainty is a commonly unacknowledged source of error. The GIS assumes that locations are exact and correct, but sometimes there is enough error in the sample point location that the actual grid cell value may not be the best classification value. This extension offers a method to estimate classification values based on a circular neighborhood around the sample point if you feel there is significant uncertainty about the location.

Sample Point
Pine
Oak
Mixed Conifer
Grassland

Sample Point located on Grassland cell (G)



Cell Areas within Circle

Cell A:	Class = Oak	Area = 0.060 hectares
Cell B:	Class = Pine	Area = 0.838 hectares
Cell C:	Class = Pine	Area = 0.911 hectares
Cell D:	Class = Mixed Conifer	Area = 0.133 hectares
Cell E:	Class = Pine	Area = 0.706 hectares
Cell F:	Class = Pine	Area = 1.517 hectares
Cell G:	Class = Grassland	Area = 1.517 hectares
Cell H:	Class = Grassland	Area = 0.917 hectares
Cell I:	Class = Pine	Area = 0.639 hectares
Cell J:	Class = Pine	Area = 1.517 hectares
Cell K:	Class = Grassland	Area = 1.517 hectares
Cell L:	Class = Grassland	Area = 0.850 hectares
Cell M:	Class = Pine	Area = 0.020 hectares
Cell N:	Class = Pine	Area = 0.644 hectares
Cell O:	Class = Grassland	Area = 0.716 hectares
Cell P:	Class = Mixed Conifer	Area = 0.065 hectares

Cell Areas per Class

Pine: 6.791 hectares
Grassland: 5.517 hectares
Mixed Conifer: 0.198 hectares
Oak: 0.060 hectares

Final Classification will be Pine

[1]**Jeff Jenness**, GIS Analyst/Wildlife Biologist
USFS Rocky Mountain Station
2500 S. Pine Knoll Drive
Flagstaff, Arizona, USA 86001
jjenness@fs.fed.us, jeffj@jennessent.com

[2]**J. Judson Wynne**
USGS - Southwest Biological Science Center
Colorado Plateau Research Station
2255 North Gemini Drive
Flagstaff, AZ 86011 USA
james.wynne@nau.edu

REFERENCES:

Boone, R.B. and W.B. Krohn. 2002. Modeling tools and accuracy assessment. In Predicting Species Occurrences: Issues of Accuracy and Scale (J. Scott, R. Heglund, M. Morrison et al.). Island Press, Washington. Pps 265-270.

Congalton, Russell; Green, Kass. (1999). Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. Lewis Publishers.

Csuti, B. and P. Crist. 1998. Methods for assessing accuracy of animal distribution maps. In Gap Analysis Program report (Cassidy, K., E.O. Garton, W.B. Krohn, L.S. Mills, J.M. Scott, K. Williams and B. Csuti, editors). University of Idaho, Moscow.

Drost, C.A., S.R. Jacobs and KA. Thomas. 1999. Accuracy assessment of vertebrate distribution models for Arizona Gap. USGS-Colorado Plateau Research Station technical report. Flagstaff, Arizona. Pps. 58.

Fielding, Alan H.; and Bell, John F. 1997. A review of methods for the assessment of prediciton errors in conservation presence/absence models. Environmental Conservation 24: 38–49.

Hess, G.R. 1994. Pattern and error in landscape ecology: a commentary. Landscape ecology 9: 3-5.

Hess, G.R. and J.M. Bay 1997. Generating confidence intervals for composition-based landscape indexes. Landscape Ecology 12: 309-320.

Luck, M. and J. Wu 2002. A gradient analysis of urban landscape pattern: a case study from the Phoenix metropolitan region, Arizona, USA. LandscapeEcology 17: 327-339.

This Kappa Analysis ArcView extension is available for free download at http://www.jennessent.com/arcview/kappa_stats.htm