NATIONAL OAK WILT SYMPOSIUM

JUNE 4-7, 2007 • AUSTIN, TEXAS

The Proceedings of the 2nd National Oak Wilt Symposium

Edited by:

Ronald F. Billings
David N. Appel

Sponsored by

International Society of Arboriculture – Texas Chapter

Cooperators

Texas Forest Service
Texas AgriLife Extension Service
The Nature Conservancy of Texas
Lady Bird Johnson Wildflower Center
USDA Forest Service, Forest Health Protection

2009

USING CLASSIFICATION TREE ANALYSIS TO PREDICT OAK WILT DISTRIBUTION IN MINNESOTA AND TEXAS

Marla C. Downing
USDA Forest Service
Forest Health Technology Enterprise Team
2150 Centre Ave., Building A, Suite 331
Fort Collins, CO 80526-1891
Email: mdowning@fs.fed.us

Vernon L. Thomas
INTECS International, Inc.
2625 Redwing Road, Suite 360
Fort Collins, CO 80526
Email: vlthomas@fs.fed.us

Jennifer Juzwik
USDA Forest Service
Northern Research Station
1561 Lindig Avenue
St. Paul, MN 55108
Email: jjuzwik@fs.fed.us

David N. Appel
Texas A&M University
Department of Plant Pathology and Microbiology
College Station, TX 77843
Email: appel@ag.tamu.edu

Robin M. Reich
Department of Forest Sciences
Colorado State University
Fort Collins, CO 80523
Email: robin@cnr.colostate.edu

and

Kim Camilli
Texas Forest Service, Forest Pest Management
Current address: Cal Fire & Cal Poly
4050 Branch Rd.
Paso Robles, CA 93446
Email: k_camilli@yahoo.com

ABSTRACT

We developed a methodology and compared results for predicting the potential distribution of *Ceratocystis fagacearum* (causal agent of oak wilt), in both Anoka County, MN, and Fort Hood, TX. The Potential Distribution of Oak Wilt (PDOW) utilizes a binary classification tree statistical technique that incorporates: geographical information systems (GIS); field sample data; commonly available, inexpensive, coarse-resolution auxiliary data; and satellite imagery from both Landsat Thematic Mapper (TM) and SPOT to predict the spatial distribution of oak wilt. Two types of model evaluations were conducted - a ten-fold cross validation and an assessment using additional oak wilt data that had been verified in the field. These evaluations indicated that at the landscape scale PDOW correctly models the presence of oak wilt, and accurately predicts oak wilt distribution in Anoka County, MN and Fort Hood, TX. Variables that were common for predicting oak wilt distribution in both Anoka County and Fort Hood were: Landsat TM Bands 3, 5, and 7; sand; aspect; and elevation. Additional variables important in Fort Hood included: Spot band 1, stream density, and slope. Variables that were unique and important for Anoka County included: TM Band 4, organic matter, silt, drainage, population, and population change.

Key words: *Ceratocystis fagacearum,* classification tree, decision tree, Landsat TM, spatial statistics, SPOT satellite imagery.

The oak wilt fungus, *Ceratocystis fagacearum* (Bretz) Hunt, kills thousands of oak trees (*Quercus* spp.) annually within the US disease range, i.e. 22 eastern states and Texas (Appel and Maggio 1984, Juzwik 2000, O'Brien et al. 2003). Continued spread of the pathogen results in expanding disease foci and establishment of new foci in forests, woodlots, and home landscapes. In Minnesota, oak wilt occurs within 20 counties. Between the years of 1991 - 2001, 6,976 acres were treated of a total estimated 15,359 acres affected (MNDNR 2001). The disease is most severe on deep sand soils in east-central Minnesota. In a one year example, the Minnesota Department of Natural Resources (MNDNR) identified and treated 3,182 acres of infected oak wilt trees in 1998 within Anoka County. The MNDNR projected that, at that infection rate, there would be a two-fold increase in oak wilt by 2008 (MNDNR 2000).

In Texas, in 2007, oak wilt occurred within 60 counties and was estimated to affect a minimum of 6,500 acres (Texas Forest Service 2007). The disease is particularly severe on the Edwards Plateau of central Texas. Within this region, including the Fort Hood military installation, the live oak / Ashe juniper community type is critical habitat for two rare bird species (the golden cheeked warbler and the black-capped vireo) indigenous to the region (Diamond 1997). Oak wilt is a potential threat to live oak in this critical habitat.

Aerial and ground surveys are regularly conducted in both states to detect new disease centers and estimate area of land affected. The recent development of new geospatial techniques and geo-statistical analysis tools offer new methods for displaying oak wilt distribution, predicting disease occurrence in areas where data is lacking, and obtaining estimates of both land area affected and forest areas at risk to the disease.

Statistical techniques such as decision tree models are useful for classification problems where mixes of both continuous and categorical data are available for geospatial analysis. Classification or decision trees are non-linear tests made up of a collection of rules displayed in the form of a binary tree. The rules are determined by a recursive partitioning procedure (MathSoft 1999). Advantages of using decision trees include the non-parametric nature of the model, ease of interpretation, and the robustness of the test (De'Ath and Fabricius 2000). Classification trees offer a way to describe the spatial continuity that is an essential feature of many natural phenomena (Isaaks and Srivastasa 1989), and have been used to: classify remote sensing imagery (Friedl and Brody 1997, Michaelson et al. 1994, Joy, Reich and Reynolds 2003), predict spatial patterns and develop indicators of hemlock woolly adelgid infestation, (Koch 2005), model *Phythophthora ramorum* (sudden oak death) distribution in California (Kelly and Meentemeyer 2002), model the presence and absence of lichen and past fires in Jalisco, Mexico (Reich, Aguiree-Bravo and Bravo et. al. 2005), and to estimate fuel loads in the Black Hills, SD (Reich, Lundquist and Bravo 2004).

The objective of our study was to determine the feasibility of using a decision tree with commonly acquired spatial datasets and location data collected in Anoka County and at Fort Hood to develop spatially-explicit maps that predict the distribution of oak wilt in these locations. Our analysis utilized field data, remotely sensed satellite data, as well as other spatial data within a geographical information system (GIS). The resulting models were evaluated for their accuracies in predicting presence or absence of the oak wilt disease.

MATERIALS AND METHODS

Study Areas

Anoka County is 110,000 hectares located in east-central Minnesota and occurs largely within the Anoka Sand Plains ecological subsection (MNDNR 1999). The terrain is a broad, flat, sandy plain with gently rolling topography. Soils are largely well-drained fine sands. The vegetation included species associated with oak openings and oak barrens. The predominant oak species are northern pin oak (*Q. ellipsoidalis*), northern red oak (*Q. rubra*) and bur oak (*Q. macrocarpa*).

Fort Hood is approximately 87,900 hectares in size, (Ribanszky and Zhang 1992). Fort Hood is located in Bell and Coryell counties, TX, within the Crosstimbers and Southern Tallgrass Prairie and the northeastern edge of the Edwards Plateau Ecoregions. Vegetation in the area consists mainly of open grasslands or savannah with individuals or mottes of oak (*Quercus* spp); ashe juniper (*Juniperus ashei*); and mixed forest dominated by oak-juniper (Diamond 1997, The Nature Conservancy 1997).

Location and Spatial Data

Presence/Absence (Dependent Variable). For the Anoka County study area, a Dependent Variable GIS Sample Point Theme was created using the Land Management Information Center (LMIC) oak wilt database as our primary data source (Table 1). Many sample locations were acquired from the 1998 LMIC oak wilt "treated" polygon data. Additional LMIC sample locations, coded as "possible active" oak wilt sites during the 1998 growing season, were randomly selected and visited in July and August, 2002. If evidence suggested the sites actually had active oak wilt infection centers in 1998 then GPS (Garmin E-Trex Legend) coordinate system points were collected for the oak wilt positive tree locations. Healthy oak site locations were also acquired during the 1998 growing season and again in September 2004. Of the 489 sample points collected in Anoka County, 156 were identified as being healthy oak sample points, and 333 were identified as having been active oak wilt sites in 1998.

All polygon centroid locations from the LMIC database and our additional sample point locations were merged to create the final dependent variable GIS Sample Point Theme. Healthy oak wilt sample point locations were assigned a value equal to 1, and oak wilt sample point locations were assigned a value equal to 2. Polygon centroid location points were acquired to create sample points from the polygon (USDA Forest Service, FHTET 2007a).

All Fort Hood dependent variable plot data for oak wilt and non-oak wilt sites were collected in the field during the growing seasons of 2003 – 2004 (Table 1). A systematic cluster plot sampling design was implemented to attain the dependent variable sample data; the ratio was two healthy plots to one oak wilt plot. The cluster plots were configured such that four 10 m x 10 m secondary sampling units (ssu) composed one 20 m x 20 m primary sampling unit (psu). This sampling design was used to avoid periodicity in the resource and to permit plots to occur at random distances for spatial modeling (Reich, Aguirre-Bravo and Bravo 2005).

Independent Variable Data. Twenty-three auxiliary or independent grid themes were constructed for use as independent variables in the Anoka County analysis (Table 2): fourteen were created from two, multi-temporal Landsat 5 TM data sets (May and September 1998). The other nine variables were: aspect, distance-to-lakes, distance-to-streams, drainage, elevation, landform, road density, slope, and stream density.

Twenty four independent variable grid themes were constructed for the Fort Hood analysis (Table 2). Seven were from Landsat 5 TM, another four were from SPOT 5 satellite imagery. The remaining thirteen variables included: aspect, detritus, elevation, forb percent, land cover, landform, organic matter, road density, slope, sand, silt, clay, and stream density. All variables for each study area were collected, aggregated or re-sampled to a 30 m x 30 m spatial resolution.

Stratification of Anoka County Land Area

The southern section of Anoka County has a higher degree of urban coverage than the northern section of the county. To determine whether spatial correlation exists between oak wilt and urban or natural landscape features, and to ensure that the urban condition in the south was not affecting the results of the model for the non-urban area to the north, the county was stratified into urban and non-urban datasets and two models were created.

Spatial Information Databases

Three spatial information databases were created for the study areas: Anoka – urban; Anoka – non-urban; and Fort Hood. To do this, information was extracted from each of the independent variable data themes at the grid cell location coincident with the sample point (Anoka County), or cluster plot (Fort Hood) locations (USDA Forest Service, FHTET, 2007b).

Classification Tree Analyses: Creation of a Disease Map

The Spatial Information Databases were used for the classification tree analyses to predict the distribution of oak wilt. The output from the classification tree was the input for conditional statements (ESRI CON statements, 2000), which were used to create an oak wilt presence or absence raster grid surface for each study area (Figs. 1A and B). Grid theme cells with values of 1 indicated lower probabilities of oak wilt presence (defined as absence). Grid theme cells with values of 2 indicated higher probabilities of oak wilt presence.

Evaluations

There were two evaluations performed in each study area: 1) the initial evaluation estimated as a sample-based misclassification error rate, and 2) the tenfold cross-validation, (Efron and Tibshirani 1993), calculated in S-PLUS© as part of the classification tree procedure.

The sample based misclassification error evaluation was conducted by intersecting oak wilt points and polygons with each of the final surfaces to determine the rate at which we accurately predicted the presence of oak wilt. For the Anoka County urban model, a total of 164 known oak wilt polygons, with a mean size of 0.76, minimum size of 0.07, and a maximum size of 10.01 acres, were used. In the Anoka County non-urban model, a total of 65 known oak wilt polygons, with a mean size of 1.94, minimum size of 0.14, and a maximum size of 13.16 acres, were used. In Anoka County, the predicted PDOW was quantified for three categories; 50, 75, and 100 percent of the assessment polygon. To quantify the number of polygons successfully predicted with oak wilt, the assessment polygon was intersected with the results from the PDOW surface, and then the area of predicted oak wilt within the assessment polygon was divided by the total area of the assessment polygon. The polygons that were accurately predicted as having oak wilt were totaled within each category (i.e., 50, 75, and 100). The total number of polygons from each category was then divided by the total number of assessment polygons used for an overall estimate of accuracy.

The cross-validation procedure validates the tree sequence by shrinking and/or pruning the tree by portioning the data into a number of subsets, fitting sub-tree sequences to these, and using a subset previously held out to evaluate the sequence. This procedure was used to identify the tree size that minimized the prediction error.

RESULTS

Oak Wilt Distribution: Models and Surfaces

Anoka County. There were thirteen terminal end nodes in the urban model, which accounted for 84 percent of the variability. The independent variables important in predicting the presence or absence of oak wilt in the urban model were: sand, TM band 4 (May 1998), TM band 6 (September 1998), aspect, silt, drainage, elevation, and population (Fig. 2A). The Anoka County non-urban model had twelve terminal end nodes, which accounted for 86 percent of the variability in the model. The independent variables important for predicting the presence or absence of oak wilt in the non-urban model were: TM band 3 (May 1998), population change, organic matter, silt, TM band 3 (September, 1998), TM band 7 (May 1998), landform, and TM band 5 (May 1998), (Fig. 2B). Oak wilt is predicted to be present in 51 and 56 percent of the urban and non-urban forests, respectively, of Anoka County (Table 3; Figs. 2A and 2 B).

Fort Hood. There were sixteen terminal nodes, which accounted for 93 percent of the variability in the Fort Hood model. The variables of importance for predicting the potential of oak wilt were: landcover, TM band 7 (May 16, 2003), SPOT 5 band 1 (July 2003), stream density, aspect, TM bands 3 and 5 (May 2003), slope, sand and elevation (Fig. 2C). Oak wilt is predicted to be present on 41 percent of the forest land at Fort Hood (Table 3; Fig. 2C).

Model Evaluations

The classification tree selected through cross validation for each of the Anoka County urban and non-urban models had misclassification errors of 0.1588 and 0.1445, respectively. The misclassification error rate for Fort Hood was 0.0651. The error matrix (Table 4) showed the following for each model:

* Anoka urban model, 83 percent (n = 47) of non-oak wilt sample points and 84 percent (n = 186) of oak wilt sample points were correctly classified,

* Anoka non-urban model, 88 percent (n = 109) of non-oak wilt sample points and 84 percent (n = 147) of oak wilt sample points were correctly classified,
* Fort Hood model, 97 percent (n = 278) of healthy sample points and 83 percent (n = 106) of oak wilt sample points were correctly classified.

A second evaluation was conducted on the oak wilt predictions for each study area using the additional data collected in the field, including 1999 and 2000 LMIC data for Anoka County, These accuracy assessment points and polygons (= test data) were not part of the dataset used to develop the oak wilt models.

The frequency accuracies for the predicted presence of oak wilt in the assessment polygons in Anoka County depended on the proportion of the area of each polygon considered. Accuracies were highest when 50% of the assessment polygons were predicted to have the disease (93 percent urban model; 88 percent non-urban model) and lowest when 100 percent of the polygon were predicted to have oak wilt (70 percent, urban model; 25 percent, non-urban model). Of 34 points known to have oak wilt in the Texas area, 25 (74 percent) were predicted to have the disease using the Fort Hood model.

DISCUSSION AND CONCLUSIONS

We restricted our analyses to variables that were easily obtained and at a minimum cost. We showed that using a classification tree on commonly-acquired datasets could reliably predict the distribution of oak wilt in Anoka County, MN, and Fort Hood, TX. The classification tree technique identified several independent variables that were useful in predicting the potential distribution of oak wilt in the Minnesota and Texas landscapes. The combinations of variables included higher values of Landsat TM Band 3 combined with low organic matter, higher values of Landsat TM Band 7 where there were more streams, and were either low lying or with flattened slopes.

Remotely-sensed satellite data, combined with location data collected in the field, was useful for identifying the presence and/or absence of oak wilt in each study area. The satellite bands selected by the classification tree were Landsat TM Bands 3, 4, 5, 6, and 7 and SPOT Band 1. Since SPOT Band 1 provides similar spectral information to the TM Bands, and the improved spatial resolution was not required, we believe there is no added benefit to including the SPOT data in the future.

Also, just as a western aspect was identified by Bowen and Merrill (1982) as being important in predicting oak wilt in Pennsylvania, aspect was identified as an important variable in modeling oak wilt in both the Anoka County urban model and the Fort Hood model. Although it is possible that aspect may be representing flat terrain.

Stratification by land use, specifically the urban and non-urban condition, did not indicate that land use was an important variable in modeling oak wilt. Future analyses should include the urban and non-urban condition as an independent variable for predicting the potential for oak wilt.

Land cover was used in the Fort Hood model but not in either of the Anoka models. As one would expect, the most important variable for predicting oak wilt in Fort Hood was for the sample location to occur in deciduous forested land cover types.

Although our potential distribution of oak wilt might be considered a theoretical construct (Felicisimo et al. 2002), our accuracy assessment using additional oak wilt locations establishes that the classification tree analysis of large-scale, commonly-acquired data can be successfully used to construct a model for predicting the potential distribution of oak wilt in both the

Minnesota and Texas landscapes. The authors recommend the continued investigation of such techniques on other forest pest species.

ACKNOWLEDGEMENTS

The authors would like to thank the following individuals: From the USDA Forest Service, Forest Health Protection: Joseph O'Brien and Dale Starkey for their pathology and forestry expertise as well as their administrative and technical support. From the USDA Forest Service, North Central Research Station, Forest Disease Unit: Kathy Kromroy for her intellectual contributions during the conceptualization and execution of the project; Kathy Ward and Paul Castillo for their technical assistance in collecting oak wilt field data and the acquisition of appropriate data layers for use in the analysis. From the U.S. Army III Corps and Ft. Hood Program: John Cornelius (Director for Endangered Species Management) and Charles Pekins for their technical and field support. From the DPW ENV Natural Resources Branch, Fort Hood, TX: Steve Jester for site characterization. From The Nature Conservancy, Texas Chapter, Fort Hood Project: Sheila Jackson for field work. Texas A&M University, Department of Plant Pathology and Microbiology: Coy Crane and Olivie Schill for field work. From the Minnesota Department of Natural Resources, Division of Forestry: Susan Burks for her participation as a stakeholder.

LITERATURE CITED

Appel, D.N., and R.C. Maggio. 1984. Aerial survey for oak wilt incidence at three locations in central Texas. Plant Disease 68: 661-664.

Bowen K.L., and W. Merrill. 1982. Oak wilt foci related to ridge bearing and aspect in Pennsylvania. Plant Disease 66: 137-139.

De'ath, G., and K.E. Frabricus. 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis. Ecology 81: 3178-3192.

Diamond, D. D. 1997. An old-growth definition for western juniper woodlands: Texas Ashe juniper dominated or codominated communities. USDA Forest Service, Southern Research Station General Technical Report SRS-15. 10 pp.

Efron, B., and R.J. Tibshirani. 1993. An introduction to the bootstrap. Monographs on Statistics and Applied Probability, Volume 57. Chapman and Hall, New York, NY. 456 pp.

ESRI, Environmental Services Resources Inc. 1992-2000. ArcView GIS 3.2a. Redlands, CA.

Felicisimo, A.M., E. Frances, J.M. Fernandez, A. Gonzalez-Diez, and J. Vara. 2002. Modeling the Potential Distribution of Forests with a GIS. Photogrammetric Engineering & Remote Sensing 68(5): 455-461.

Friedl, M.A., and C.E. Brodley 1997. Decision tree classification of land cover from remotely sensed data. Remote Sensing of Environment 61: 399-409.

Isaaks, E.H., and R.M. Srivastava. 1989. An Introduction to Applied Geostatistics. Oxford: Oxford University Press. 561 pp.

Juzwik, J. 2000. An oak wilt primer. The Journal of the International Oak Society 11: 14- 20.

Joy, S.M., R.M. Reich, and R.T. Reynolds. 2003. A non-parametric, supervised classification of vegetation types on the Kaibab National Forest using decision trees. International Journal of Remote Sensing 24(9): 1835-1852.

Kelly, M., and R.K. Meentemeyer. 2002. Landscape dynamics of the spread of sudden oak death. Photogrammetric Engineering and Remote Sensing 68(10): 1001-1011.

Koch, F. H., H.M. Cheshire, and H.A. Devine. 2005. Mapping hemlocks via tree-based classification of satellite imagery and environmental data, Pp. 104-114. *In*: Brad Onken and Richard Reardon (eds.), The Third Symposium on Hemlock Woolly Adelgid in the Eastern United States. USDA Forest Service, Forest Health Technology Enterprise Team. FHTET-2005-01.

MathSoft, 1999. S-PLUS 2000: Modern Statistics and Advanced Graphics; Guide to Statistics, Volume 1: 369-401.

McNab, H.W. 1989. Terrain shape index: Quantifying effect of minor landforms on tree height. Forest Science. 35(1): 91-104.

Michaelson, J., D.S. Schmiel, M.A. Friedal, F.W. Davis, and R.O. Dubayah. 1994. Regression tree analysis of satellite and terrain data to guide vegetation sampling and surveys. Journal of Vegetation Sciences 5: 673-686

MNDNR, Minnesota Department of Natural Resources Division of Forestry. 1999. Ecological Subsections of Minnesota. [System ecssbne2]. http://deli.dnr.state.mn.us/.

MNDNR, Minnesota Department of Natural Resources Division of Forestry - MIS Bureau. 2000. Land Type Associations of Minnesota. [System ecsltpy2]. http://deli.dnr.state.mn.us/.

MNDNR, Minnesota Department of Natural Resources Division of Forestry, Forest Insect and Disease Newsletter, June 29, 2001. http://www.dnr.state.mn.us/fid/july01/feature.html#table

O'Brien, J.G., M.E. Mielke, D. Starkey, and J Juzwik. 2003. How to identify, prevent, and control oak wilt. USDA Forest Service, Northeastern Area State and Private Forestry. NA-PR-03-00. http://www.na.fs.fed.us/spfo/pubs/howtos/ht_oakwilt/toc.htm. (August 28, 2003).

Reich, R.M., C. Aguirrie-Bravo, and V.A. Bravo. 2005. Analysis of statistical strategies for advancing monitoring and assessment of terrestrial ecosystems in the Mexican state of Jalisco: Pilot study synthesis. Submitted to: Fideicomiso para la Administracion del Programa de Desarrollo Forestal del Estado de Jalisco, Guadalajara, Jalisco, Mexico. 136 pp.

Reich, R.M., J.E. Lundquist, and V.A. Bravo. 2004. Spatial models for estimating fuel loads in the Black Hills, South Dakota. International Journal of Wildland Fire 13: 119-129.

Texas Forest Service News Release, July 31, 2007. http://texasoakwilt.org/Documents/Press/Help%20stop%20spread%20of%20deadly%20oak%20wilt.pdf

Texas Forest Service. 2007. Occurence of tree mortality in central Texas counties, July 2007. http://www.texasoakwilt.org/Gallery/maps/Intensity_Map.jpg

USDA Forest Service, Forest Health Protection, Forest Health Technology Enterprise Team (FHTET). 2007a. Spatial Statistics: Sample Point and Polygon Centroid Generator. http://www.fs.fed.us/foresthealth/technology/spatialstatistics/index.shtml

USDA Forest Service, Forest Health Protection, Forest Health Technology Enterprise Team (FHTET). 2007b. Spatial statistics: Methods and analysis: Get model data. http://www.fs.fed.us/foresthealth/technology/spatialstatistics/index.shtml

Table 1. Each sample point dataset containing the presence and absence oak wilt data for each study area: Anoka County urban, Anoka County non-urban, and Fort Hood became the dependent variable in a classification tree against which the independent variables were tested for correlation:

Anoka County Dependent Variable Sample Point Theme
A. Oak wilt presence and absence field sample data from the Land Management Information Center (LMIC), Forest Health, Oak Wilt, treated site polygon data, 1998.
B. Field visits by the USDA FS NCRS Forest Disease Unit and the Forest Health Technology Enterprise Team (FHTET) to the LMIC, Forest Health, “active” oak wilt sites.
C. Field visits by the USDA FS NCRS Forest Disease Unit and FHTET to randomly selected healthy oak forest sites.

Fort Hood Dependent Variable Sample Point Theme
A. Field visits by the USDA FS Forest Health Technology Enterprise Team (FHTET), Region 8, Texas A & M and Texas Department of Forestry personnel, to randomly select “healthy and “active” oak wilt sites.

Table 2. All independent variables used to determine the level of correlation with the dependent variable in the binary classification trees. Anoka County, MN: A-K; Fort Hood, TX: L-Z.

Anoka County Independent Variables (A-K)

A.	Aspect (compass direction), derived from the USGS DEM using ArcView Spatial Analyst (ESRI) aspect function; North = A, Northeast = B, East =C, Southeast = D, South =E, Southwest = F, West = G, Northwest = H.
B.	Distance to Lakes: USGS 1:100,000 DLG data, measured using ArcView Spatial Analyst, distance in meters from feature function.
C.	Distance to Streams: USGS 1:100,000 DLG data, measured using ArcView Spatial Analyst, distance in meters from line feature function.
D.	Drainage: from USDA, NRCS, SSURGO Data Version 2.1 (December 2003); http://soildatamart.nrcs.usda.gov
E.	Elevation: derived from the USGS 30 *m* (1:24000) DEM.
F.	Landform: (independent of slope), created from a custom ArcView Avenue application, which uses an irregular 3 x 3 kernel, where positive values indicate concavity and negative values indicate convexity, to calculate landform from a USGS DEM. A zero value indicates flat terrain (McNab, 1989).
G.	Seven Bands from Landsat Thematic Mapper Satellite Imagery: Path 27 Row 29 bands 1-7, acquired May 1998.
H.	Seven Bands from Landsat Thematic Mapper Satellite Imagery: Path 27 Row 29 bands 1-7, acquired September 1998.
I.	Road Density: measured using ArcView Spatial Analyst, distance in meters from line feature function. It was calculated as the sum of roads within 400 x 400 meter grid surfaces. Roads include City Streets, County Roads, and TWP Roads from USGS 1:24,000 data and Major and Ramp roads from MN Department of Transportation data.
J.	Slope degrees: derived from the USGS DEM using ArcView Spatial Analyst (ESRI) slope function
K.	Stream Density: from Minnesota Department of Natural Resources, MN Wetlands and Surface Water Resources data set; calculated as the sum of all stream surface area within 400 x 400 meter surface grids.

Fort Hood Independent Variables (L-Z)

L.	Aspect (compass direction), derived from the USGS DEM using ArcView Spatial Analyst (ESRI) aspect function; North = 1, Northeast = 2, East = 3, Southeast = 4, South = 5, Southwest = 6, West = 7, Northwest = 8.
M.	Clay: from USDA, NRCS, SSURGO Data Version 2.1 (December 2003); http://soildatamart.nrcs.usda.gov
N.	Detritus: from USDA, NRCS, SSURGO Data Version 2.1 (December 2003); http://soildatamart.nrcs.usda.gov
O.	Elevation derived from the USGS 30 meter resolution DEM (1:24000 scale).
P.	Forb percent: from USDA, NRCS, SSURGO Data Version 2.1 (December 2003); http://soildatamart.nrcs.usda.gov
Q.	Landcover: Derived from Landsat Imagery collected 5/16/2003; a = open, b = coniferous, c = deciduous.
R.	Landform: (independent of slope), created from a custom ArcView Avenue application, which uses an irregular 3 x 3 kernel, where positive values indicate concavity and negative values indicate convexity, to calculate landform from a USGS DEM. A zero value indicates flat terrain (McNab, 1989).
S.	Seven Bands from Landsat 5 TM Satellite Imagery: Path 27 Row 38, bands 1-7, acquired 5/16/2003.
T.	Organic matter: from USDA, NRCS, SSURGO Data Version 2.1 (December 2003); http://soildatamart.nrcs.usda.gov

U.	Road Density: The Nature Conservancy, Fort Hood Project, Fort Hood, Texas.
V.	Slope degrees: derived from the USGS DEM using ArcView Spatial Analyst (ESRI) slope function.
W.	Sand: from USDA, NRCS, SSURGO Data Version 2.1 (December 2003); http://soildatamart.nrcs.usda.gov
X.	Silt: from USDA, NRCS, SSURGO Data Version 2.1 (December 2003); http://soildatamart.nrcs.usda.gov
Y.	SPOT 5 Satellite: Multi-spectral Imagery: Bands 1-4 acquired 7/29/ 2003
Z.	Stream Density: The Nature Conservancy, Fort Hood Project, Fort Hood, Texas; calculated as the sum of all stream surface area within 400 x 400 meter surface grids.

Table 3. Proportion of forest predicted with oak wilt and healthy oak in the urban and non-urban models for the Anoka County and Fort Hood study areas.

	Urban	Non-Urban	Anoka County Totals	Fort Hood Totals
Forested Area	26,179 acres	54,889 acres	81,068 acres	107,665 acres
Oak Wilt	13,452 acres 51 percent	31,014 acres 56 percent	44,466 acres 55 percent	44,192 41 percent
Healthy Oak	12,727 acres 48 percent	23,875 acres 43 percent	36,602 acres 45 percent	63,473 59 percent

Table 4. Misclassification error rates for each study area:

A. Non-Urban	Absent	Present
Classified Absent	96	24
Classified Present	13	123

B. Urban	Absent	Present
Classified Absent	39	29
Classified Present	8	157

A. Fort Hood	Absent	Present
Classified Absent	271	18
Classified Present	7	88

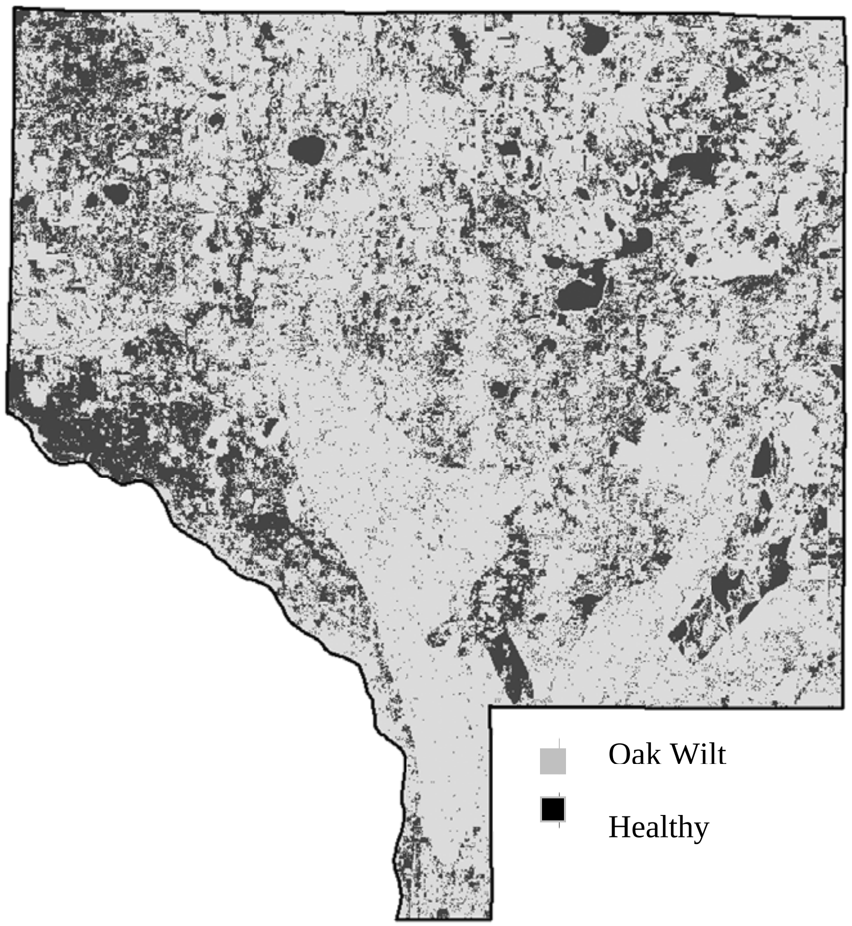


Figure 1A. Potential distribution of oak wilt with healthy oak in Anoka County, MN. The urban and non-urban models for Anoka County were merged into a single potential distribution of oak wilt grid surface according to the predicted binary output for the county.

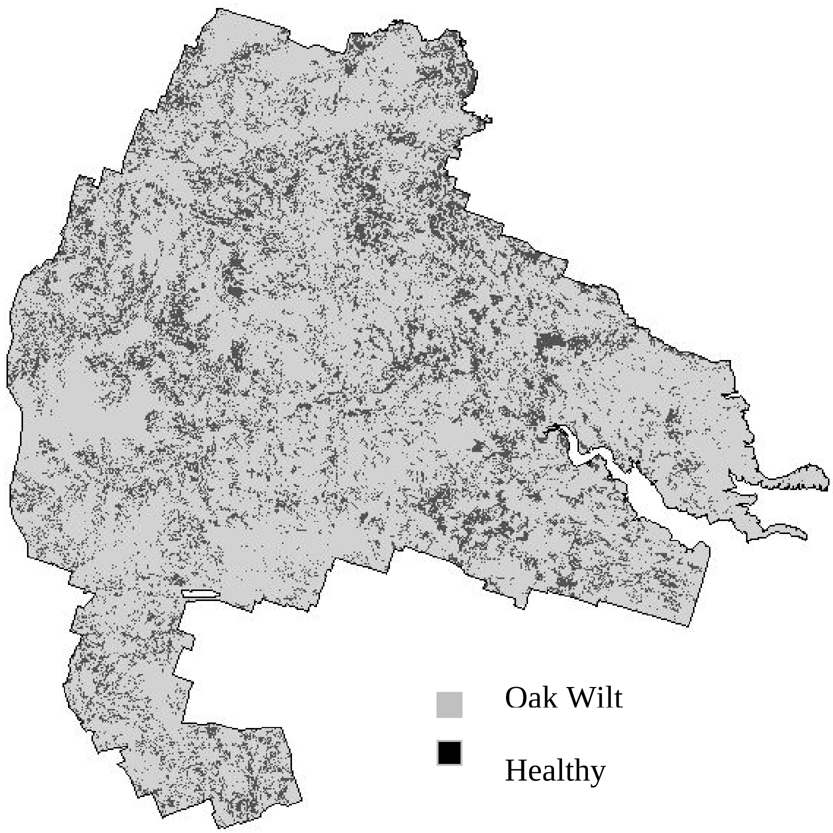
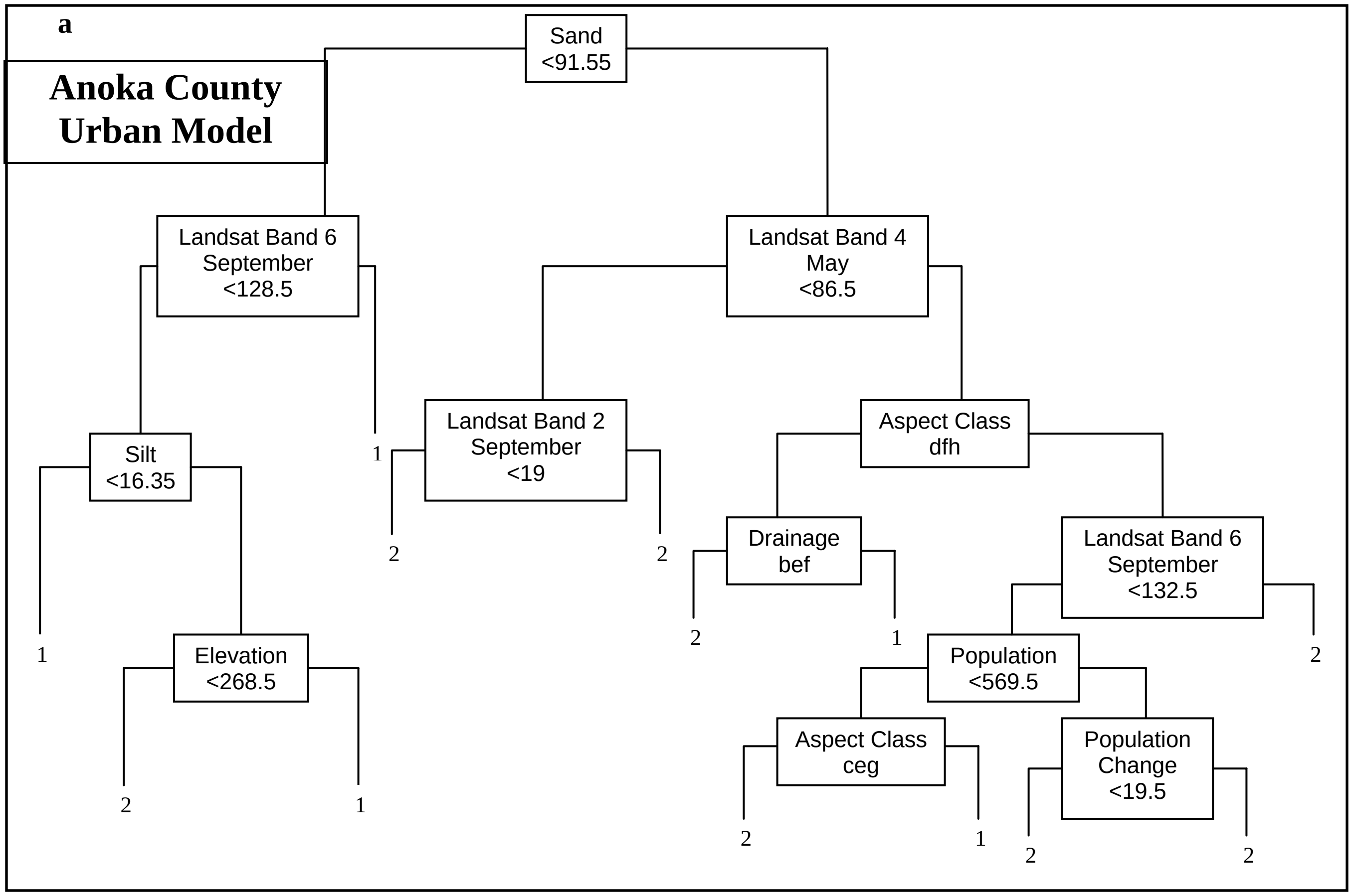
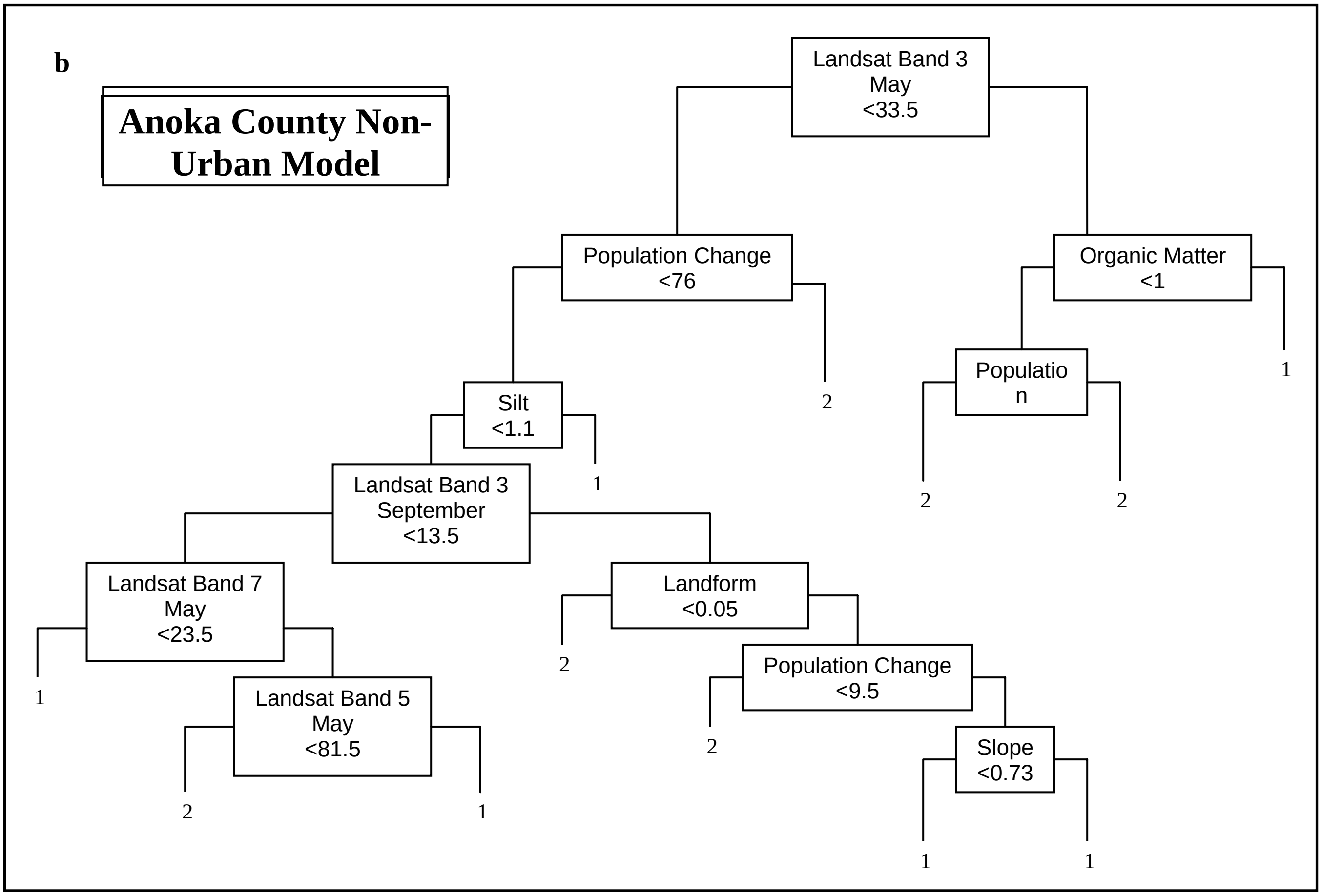


Figure 1B. Potential distribution of oak wilt and healthy oak in Fort Hood, TX.


a
Anoka County Urban Model
Sand <91.55
Landsat Band 6 September <128.5
Landsat Band 4 May <86.5
Silt <16.35
1
Landsat Band 2 September <19
Aspect Class dfh
2
2
Drainage bef
Landsat Band 6 September <132.5
1
Elevation <268.5
2
1
Population <569.5
2
2
1
Aspect Class ceg
Population Change <19.5
2
1
2
2


b
Anoka County Non-Urban Model
Landsat Band 3 May <33.5
Population Change <76
Organic Matter <1
Silt <1.1
2
Populatio n
1
Landsat Band 3 September <13.5
1
2
2
Landsat Band 7 May <23.5
Landform <0.05
Population Change <9.5
2
1
Landsat Band 5 May <81.5
2
Slope <0.73
2
1
1
1

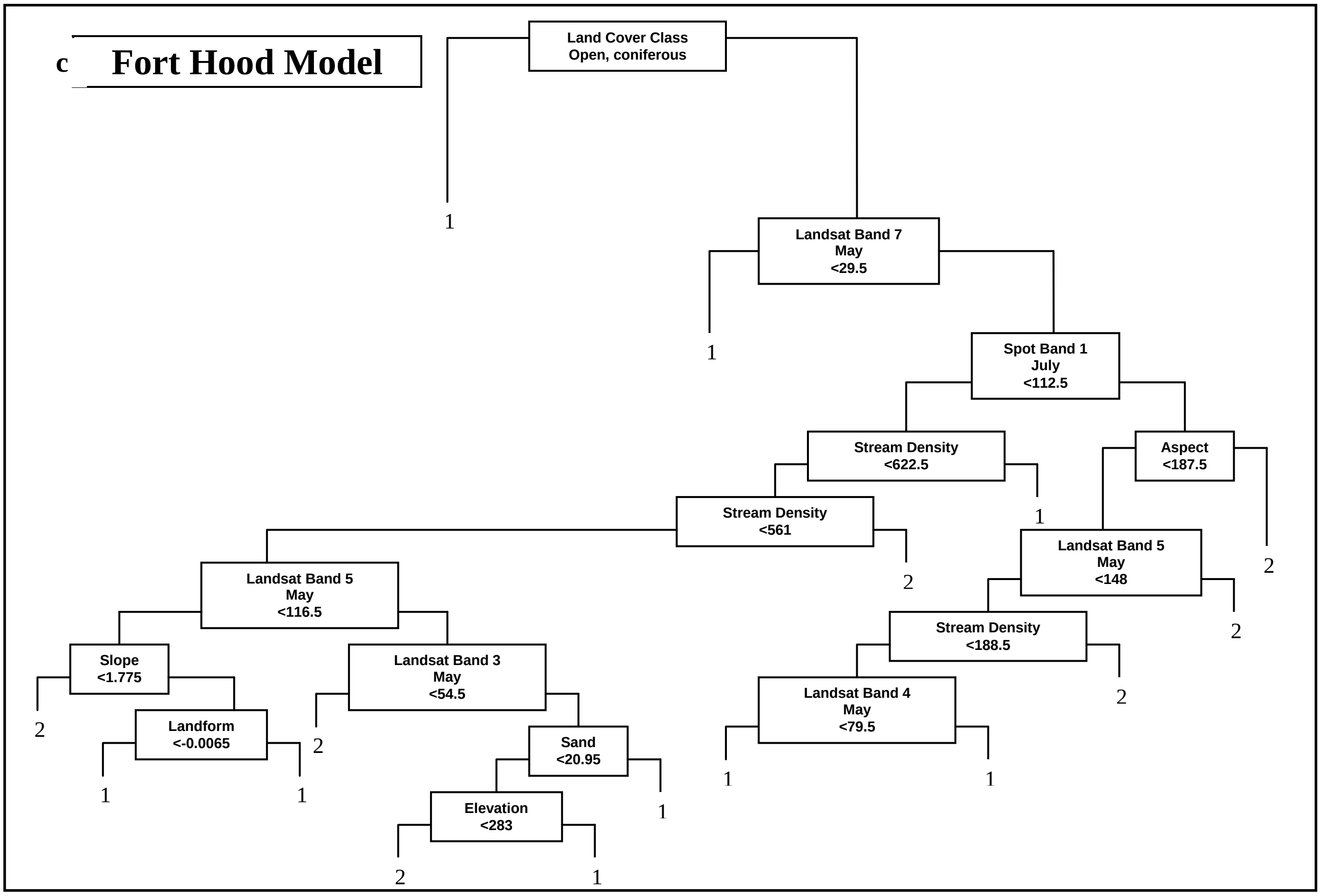


Figure 2. S-Plus classification tree output for: a) Anoka County urban model, b) Anoka County non-urban model, and c) Fort Hood model.