Informal Comparative Review of *Phytophthora ramorum* Models

DRAFT FOR REVIEW BY PARTICIPANTS OR CONFERENCE DELEGATES

Summary

Eight national-scale *P. ramorum* risk models were sub-divided into two groups: those that used *P. ramorum* observations, and those that do not. In general, models that use *P. ramorum* observations appear to predict a more restricted distribution than those that do not. A comparison of the models reveals a consistent national pattern of high risk in the coastal North-West and the central Appalachian mountains, and low risk through the Great Plains. The models differ in predicting the extent of risk in the Northeast, coastal California, parts of the Southeast and the Northern Midwest. The greatest uncertainty appears to be: i) the extent of the Northern boundary of risk east of the Rockies; and ii) the degree of risk in the Southeast. Comparing the output from multiple models gives decision makers more information than a single model alone. In particular, where multiple models agree, decision makers can have greater confidence in the results than where models disagree.

Participants

Roger Magarey, USDA APHIS CPHST/North Carolina State University (Editor)
Betsy Randall-Schadel, USDA APHIS CPHST (Editor)
Glenn Fowler, USDA APHIS CPHST
Maggi Kelly, University of California - Berkeley
Frank Koch, USDA Forest Service
Dan McKenney, Canadian Forest Service
Ross Meentemeyer, University of North Carolina-Charlotte, formerly at Sonoma State University
Steve Miller, Canada Food Inspection Agency
Bill Smith, USDA Forest Service
Robert Venette, USDA Forest Service, North Central Research Station

Objective

There have been numerous efforts to create predictive models for *Phytophthora ramorum,* causal agent of Sudden Oak Death. The purpose of this report is to review prediction models for *P. ramorum* which have been developed within the United States and Canada. The scope of this review has been limited to national-scale risk models, although several other studies at the regional or state scale exist. The predictive models reviewed have been created by a variety of methods including rule-based relationships, infection models, training algorithms based on presence and absence data, and a combination of methods. Since different modeling approaches may create different recommendations for survey, detection and eradication, APHIS Emergency and Domestic Programs requested a comparative review of models, particularly with respect to northern climates. To facilitate this task, an APHIS-CPHST organized a conference call with participants from the United

States and Canada in September, 2005. This call was followed by a meeting held in early November in Asheville, NC sponsored by the US Forest Service.

Model Descriptions

The most important sub-division of the predictive models is their need for *P. ramorum* observations (distribution). Models that do not use *P. ramorum* observations use rules or models to predict pathogen distribution based on climate or other variables. (Smith, 2002; Smith et al., 2002; Kelly et al., 2005; Mentemeyer et al., 2004; Magarey et al., 2005a; Venette and Cohen 2006). Other models use rules or algorithms to predict *P. ramorum* distribution based on relationships between the observed pathogen distribution and climate or other variables (McKenney et al., 2003, Kelly et al., 2005).   Models that do not use observed pathogen data may oversimplify relationships with climate variables. For example, the NAPPFAST system (Borchert and Magarey, 2004) uses an infection model (Magarey et al., 2005b) combined with host distribution data; consequently, it may not reflect establishment. Including lethal cold temperature improves the prediction of establishment but there is some uncertainty about the lethal cold threshold since the data set (DEFRA, 2005) used to determine the lethal cold temperatures is quite limited. This leads to uncertainty in the northern boundary of risk in the NAPPFAST model. Models that use *P. ramorum* observations may underestimate disease risk due to the current restricted distribution. *P. ramorum* establishment in North America is currently confined to mild coastal climates; therefore, extrapolation based on presence/absence data may be problematic. Using a combination of models, such as the Kelly combination model, may improve this estimate (Kelly et al., 2005). Nursery observations are more widely distributed but may be questionable, since these observations do not represent establishments in ambient environment.

For each model, we compared the spatial resolution, the climate input variables, host data, requirement for *P. ramorum* observations, variables selection and outputs (Table 1). The spatial resolution varies from 1 to 10 km^2. Most models use annual or monthly climate variables (usually averages or sums); the NAPPFAST model uses daily climate data. Important variable selection is either "*A priori,*" using variables chosen by the modeler at the beginning of the analysis rather than "Through training," using variables chosen by the model that best fit the observed pathogen distribution.

Model Performance

The final model output for the United States is presented for each group (Table 2). A comparison of the models reveals a consistent national pattern of high risk in the coastal Northwest and the central Appalachian mountains, and low risk through the Great Plains. The models differ in predicting the extent of risk in the Northeast, coastal California, parts of the Southeast and the Northern Midwest. In our opinion, the greatest uncertainty is: i) the extent of the Northern boundary of risk east of the Rockies; and ii) the degree of risk in the Southeast. In general, models that use *P. ramorum* observations appear to predict a more restricted distribution than those that do not. Comparing the output from multiple models is more likely to give decision makers more information than a single model alone.

Table 1—Characteristics of selected *P. ramorum* models.

Group	Model Name	Spatial Resolution	Climate Input Variables	Host Data	*P. ramorum* Presence/ absence data required?	Important Variable Selection	Output
USFS Smith	Rule based	Unknown	Monthly max, min temperature, rainfall,	Unknown	No	*A priori*	Survey grids
CPHST NAPP-FAST	NAPPFAST	10 Km²	Daily average temperature, daily total leaf wetness hours, annual extreme min temperature	Forest cover data USFS 1991	No	*A priori*	Probability
UCB Kelly	Rule-based			Deciduous, conifer, mixed (National Land Cover Data 1992) Hardwood diversity index (Digital tree range maps for North America) Percentage forest basal area in red and live oak group (USFS N. Eastern Research Station)	No	*A priori*	Ranked
UCB Kelly	Logistic Regression		Total annual precipitation, total annual frost days, average minimum temperature, average maximum temperature and average maximum august temperature (Day Met)		Yes	Through training	Probability
UCB Kelly	Classification and Regression Tree				Yes	Through training	P/A based on # runs
UCB Kelly	Genetic Algorithm	1 km²			Yes	Through training	P/A based on # runs
UCB Kelly	Support Vector Machines				No (1-class) Yes (2-class)	Through training	P/A based on # runs
McKenneyNatural Resources Canada	Bioclimatic envelope	~10 Km² but finer resolution output possible	Annual Mean Temperature Mean Diurnal range Max, min, mean Temperature of warmest and coldest Periods Temperature Annual Range Precipitation of Warmest/coldest Quarter	Not used in this version of model	Yes Nursery data also used as a trial but not normally used They do not represent known "establishments"	Primarily through *a priori* expectations	Climatic range / suitability based on known current establishment ie other locations with same combination of climate variables
McKenney NRCan	Extreme minimum temperature model	~10 Km² but finer resolution possible	Average extreme minimum winter temperature	Not used	No	Based on *(DEFRA, 2005)*	Shows the average and variation in the –25 degree C isotherm
USFS Venette & Cohen	CLIMEX	2.5 Km²	Monthly climate normals (1971-2000): max., min. temperature, precipitation	Not used	No (Yes-Validation)	*A priori*	Ecoclimatic Index (Ranked)

Table 2. Model results

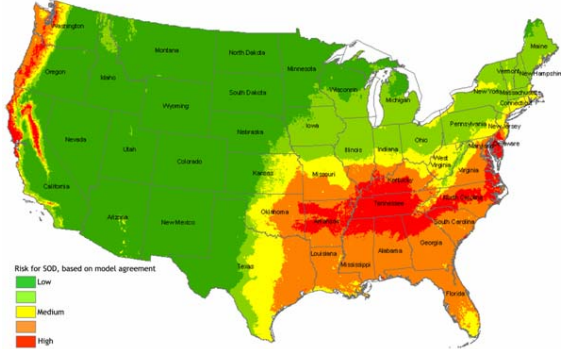| Group	Map
USFS Smith	
CPHST NAPPFAST No lethal cold exclusion (Magarey et al., 2005; Borchert and Magarey, 2005)	
CPHST NAPPFAST Lethal cold exclusion (Magarey et al., 2005; Borchert and Magarey, 2005; DEFRA, 2005)	
UC Berkeley Kelly Rule-based (Kelly et al., 2005)	

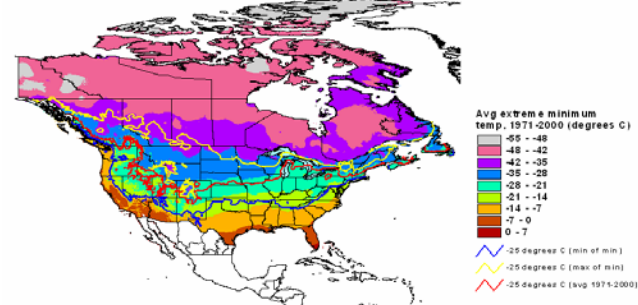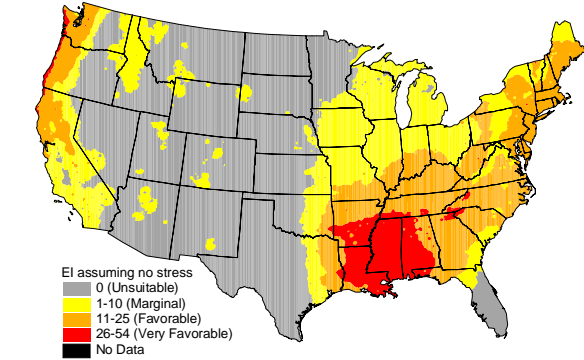**Table 2. Model results**
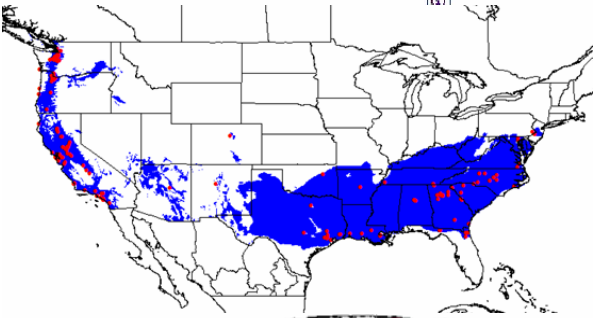
Group	Map
UC Berkeley Kelly Logistic Regression (Kelly et al., 2005)	Logistic Regression
UC Berkeley Kelly Classification and Regression Tree (Kelly et al., 2005)	CART
UC Berkeley Kelly Genetic Algorithm (Kelly et al., 2005?)	GARP
UC Berkeley Kelly Support Vector Machines (Guo et al., 2005)	SVM

Table 2. Model results

Group	Map
UC Berkeley Kelly Combination (Kelly et al., 2005)	
Natural Resources Canada *McKenney* *(Mc Kenney et al., 2003)*	
Natural Resources Canada *McKenney* *Extreme minimum* *temperatures* *(DEFRA, 2005)*	
USFS *Venette* *(Venette & Cohen, 2006)*	

References

Borchert, D. M., andMagarey, R. D. 2004. A guide to the use of NAPPFAST.
   http://www.nappfast.org

DEFRA 2005. *Phytophthora ramorum* epidemiology: infection, latency and survival.
   Department of Environment, Forestry and Rural Affairs (DEFRA) Project Report
   PH0194.

Guo, Q., Kelly, M.   and Graham, C. 2005. Support vector machines for predicting
   distribution of Sudden Oak Death in California. Ecological Modeling 182(1):75-90.

Kelly, M., Shari, Q., Guo, Q. and Liu D. 2005. Modeling risk for SOD Nationwide:
   What are effects of model choice on risk prediction? Proceedings of 2005 SOD
   Science Symposium, January 18-21, 2005. Monterey, California.

Hüberli, D., Van Sant, W., Swain, S., Davidson, J. and Garbelotto, M. 2003.
   Susceptibility of *Umbellularia californica* to *Phytophthora ramorum*. In Eighth
   International Congress of Plant Pathology, Abstract p. 268. Christchurch, New
   Zealand.

Magarey, R., Fowler, G., Randall-Schadel, B. and Colunga, M. 2005a Climate and Host
   Risk Map for Sudden Oak Death Risk (*Phytophthora ramorum).*
   http://www.nappfast.org

Magarey, R.D., Sutton, T.B. and Thayer, C. L. 2005b. A Simple Generic Infection
   Model for Foliar Fungal Plant Pathogens. Phytopathology 95:92-100.

Meentemeyer, R., Rizzo, D., Mark, W. and Lotz, E. 2004. Mapping the risk of
   establishment and spread of sudden oak death in California.   Forest Ecology and
   Management 200:195-214.

McKenney, D. W., Hopkin, A. A., Campbell, K. L., Mackey, B. G., and Footit, R.
   2003. Opportunities for improved risk assessments of exotic species in Canada using
   bioclimatic modeling. Environmental Monitoring and Assessment. 88: 445-461.

Orlikowski, L. B. and Szkuta, G. 2002. First record of *Phytophthora ramorum* in Poland.
   Phytopathologia Polonica. 25, 69-79.

Smith, W. 2002. Sudden Oak Death National Detection Survey and For Forests and
   Risk/Hazard Map. United States Forest Service.

Smith, W. D., Goheen, E. M., Sapio, F., Gottschalk, K. W. , Frankel, S.,and Dunn, P..
2002. Development of a National Survey Protocol for Detection of *Phytophthora
ramorum.* [abstract], p. 24 *in* Proceedings of the Sudden Oak Death, A Science
Symposium, USDA Forest Service and University of California, Berkeley from
http://danr.ucop.edu/ihrmp/sodsymposium.html

Venette, R.C., and Cohen, S.D. 2006. Potential climatic suitability of the eastern United
States for establishment of *Phytophthora ramorum*. Forest Ecology and Management
(Submitted)

Werres, S., Marwitz, R., Man Int Veld, W.A., De Cock, A.W.A.M., Bonants, P.J.M., De
Weerdt, M., Themann, K., Ilieva, E. and Baayen, R.P.   2001. *Phytophthora
ramorum* sp. nov., a new pathogen on *Rhododendron* and *Viburnum*. Mycol. Res.
105: 1155-1165.