

## Sudden oak death: geographic risk estimates and predictions of origins

D. A. Kluza<sup>a\*†</sup>, D. A. Vieglais<sup>b</sup>, J. K. Andreasen<sup>a</sup> and A. T. Peterson<sup>b</sup>

<sup>a</sup>U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, 8623-N, 1200 Pennsylvania Avenue Northwest, Washington, DC 20460; and <sup>b</sup>Natural History Museum and Biodiversity Research Center, University of Kansas, Lawrence, KS 66045, USA

Ecological niche modelling techniques were applied to address the questions of the origins and potential geographic extent of sudden oak death, caused by the pathogen *Phytophthora ramorum*. Based on an ecological niche model derived from the phytopathogen's California distribution and distributions of potential host species, it was determined that the disease has high potential to colonize the southeastern United States, and that its likely source area is eastern Asia.

**Keywords:** ecological niche model, genetic algorithm for rule-set prediction, *Phytophthora ramorum*, sudden oak death

### Introduction

Species of *Phytophthora* are ecologically and economically important phytopathogens that occur in natural and agricultural ecosystems (Erwin & Ribeiro, 1996). Introduced *Phytophthora* species can be particularly devastating, causing pervasive and persistent blights or rots that affect single species (e.g. Port Orford cedar root rot; *P. lateralis*; Jules *et al.*, 2002), ecological communities (e.g. *P. cinnamomi* in Australian *Eucalyptus marginata* forests; Weste, 2003) and important commercial crops (e.g. potato and tomato late blight; *P. infestans*; Fry & Goodwin, 1997).

In 2001, *P. ramorum* was described and identified as the causal agent of a twig blight affecting nursery and garden *Rhododendron* spp. in Germany and the Netherlands since 1993 (Werres *et al.*, 2001). In 2002, *P. ramorum* was identified as the agent responsible for 'sudden oak death' (SOD; Rizzo *et al.*, 2002), a syndrome first observed in 1994 and associated with widespread mortality of oaks (*Quercus* spp.) and tanoaks (*Lithocarpus densiflorus*) along the coastal ranges of California, USA. The recent emergence of *P. ramorum* and the discovery that North American and European forms are of different mating types suggest that this pathogen is an introduced species, undergoing independent introductions from its native distributional area (Werres *et al.*, 2001; Rizzo & Garbelotto, 2003).

*Phytophthora ramorum* affects a variety of species, particularly members of the Fagaceae (oaks and beeches) and Ericaceae (e.g. *Vaccinium* spp. and *Rhododendron* spp.), and hosts tend to exhibit either bark cankers or foliar symptoms (e.g. leaf lesions and twig dieback; Rizzo & Garbelotto, 2003). Among susceptible bark canker hosts (i.e. trees for which infection is usually lethal), *P. ramorum* is particularly aggressive to oaks in the section *Lobatae*, commonly referred to as red or black oaks (hereafter, red oaks; Garbelotto *et al.*, 2003; Rizzo & Garbelotto, 2003). Given the broad distribution and high diversity of North American red oaks (Jensen, 1997), the emergence of SOD has raised concerns over the possibility and consequences of this pathogen becoming established in other parts of the continent (Rizzo & Garbelotto, 2003).

To determine *P. ramorum*'s potential North American and global distribution, this study used ecological niche modelling (ENM). ENM in general involves relating known occurrence points to a suite of spatially explicit variables representing features of the ecological landscape in a predictive algorithm to produce a rule set that describes the species' ecological requirements (i.e. niche, *sensu* Grinnell, 1917). Although this technique is novel in plant pathology, it has seen extensive application in biodiversity science and human disease applications, so it represents a logical tool for application to the question of the potential geographic distribution of *P. ramorum*.

### Materials and methods

Commonly employed ENM approaches include climate matching, basic multivariate logistic regression analyses, and models based on environmental similarity or distance (Guisan & Theurillat, 2000; James & McCulloch, 2002;

\*E-mail: daniel.kluza@maf.govt.nz

†Present address: Biosecurity New Zealand, MAF, PO Box 2526, Wellington, New Zealand

Accepted 30 November 2006

Scott *et al.*, 2002; Kadmon *et al.*, 2003; Rushton *et al.*, 2004). These techniques are typically deterministic, focussing on a single decision rule or a small set of rules to describe the potential distribution of a particular species. Because distributional limits can be governed by different factors across a species' geographic range (Grinnell, 1917; Swihart *et al.*, 2003), more complex, multiple-criterion ENM approaches are desirable (Elith *et al.*, 2006). What is more, comparisons among methods are still few, and have focused on interpolation challenges, whereas the challenge at hand in this paper is one of extrapolation to continental and even global scales, so choice of methodology is not clear.

One such multiple-criterion ENM that has seen broad application in particular to extrapolative challenges (Peterson, 2003) is the Genetic Algorithm for Rule-Set Prediction (GARP), a machine-learning application. Machine-learning algorithms are computationally powerful because they are nonparametric, nonlinear and relatively unaffected by collinearity, and can explore and describe complex relationships among variables in complex solution spaces (Lek *et al.*, 1996; Olden, 2000). In the present study the decision to model the ecological niche of *P. ramorum* using GARP was made because of its multiple decision-rule capability, known robust predictive performance (Peterson & Cohoon, 1999; Stockwell & Peterson, 2002; Anderson *et al.*, 2003), and ability to project niche models onto different geographic areas.

GARP is a superset of modelling algorithms (e.g. logistic regression, range rules) that reconstructs ecological niches of species based on nonrandom associations between occurrence points and the ecological characteristics of those localities (Stockwell & Noble, 1992; Stockwell, 1999; Stockwell & Peters, 1999). Such models can be used to predict species' potential geographic distribution (Peterson *et al.*, 2002a,b; Peterson, 2003; Soberón & Peterson, 2005). Model output is spatially explicit, with potential distributions summarized in ARC/INFO raster grids. Detailed descriptions of the GARP algorithm, covered briefly below, are presented by Stockwell & Noble (1992) and Stockwell & Peters (1999).

GARP model inputs are species occurrence data (latitude/longitude) and environmental variables. Generally, a minimum of 20 occurrence points are needed to produce models that aren't overfit and unable to predict into unsampled areas (Stockwell & Peterson, 2002). Basic environmental inputs are raster geographic information system (GIS) coverages (see detailed explanations below) consisting of topographic and climate variables; biotic variables can also be included (e.g. a coverage representing host density), but typically are not used because these data are often lacking and/or difficult to represent in a spatially explicit manner. GARP partitions occurrence data into training and test datasets, the former used to build the model, and the latter used to test the accuracy of the model.

GARP describes nonrandom relationships between a species' occurrence and the environment using multiple rules. Rules are IF ... THEN statements relating to environmental conditions and whether a species is present or absent under them. GARP employs four rule types: atomic

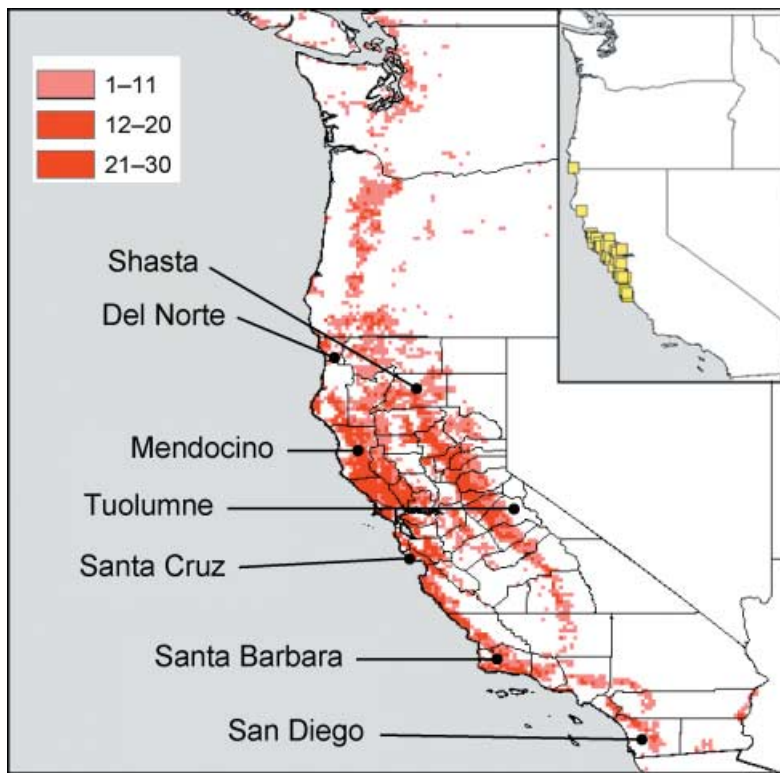
rules (use only single values of a variable), envelope rules (spanning ranges of values), negative range rules (as envelope rules, but with ranges being excluded from the prediction rather than included), and logit rules (logistic regression models) (Stockwell & Peters, 1999). After building an initial set of rules, GARP makes iterative changes to them: in each iteration ('generation'), rules undergo 'genetic' changes – insertions, deletions, point mutations, and crossing over among rules can all occur. In this manner, GARP explores solution space flexibly to identify nonrandom relationships between environmental conditions and species' occurrences. The final rule set is arrived at when iterative changes fail to improve the model's predictive accuracy, or when a user-defined number of iterations has been reached.

GARP is not a deterministic algorithm, so its output (rule sets and associated predicted distributional areas) varies from run to run. To capture the variation among potential distributions, a best-subset approach can be employed, where the best models, as determined from overprediction (commission) and underprediction (omission) errors, are selected from the final model set (Anderson *et al.*, 2003). These models can then be overlaid in a GIS to determine where they converge, that is, where spatially most or all of the models agree. GARP is not a black box – it is possible to examine the rule sets, and a jackknife feature allows correlations to be made between model performance and input variables, providing some insight into how specific variables affect predictive accuracy (Peterson & Cohoon, 1999; Peterson *et al.*, 2003a).

Locality records (116) for North American *P. ramorum* (A2 mating type; see Discussion) were acquired from the OakMapper web application (<http://kellylab.berkeley.edu/SODmonitoring/OakMapper.htm>) and the Oregon Department of Agriculture ([http://egov.oregon.gov/ODA/PLANT/path\\_sod\\_index.shtml](http://egov.oregon.gov/ODA/PLANT/path_sod_index.shtml)).

Because GARP models distributions based on an organism's presence/absence within a landscape pixel (in this analysis, pixels were  $0.08 \times 0.08^\circ$ ; see below), one point was randomly selected from each pixel where *P. ramorum* occurred, yielding 68 spatially unique localities (Fig. 1a). From these records, a training dataset was compiled by randomly selecting a single point from each  $0.25 \times 0.25^\circ$  block ( $\sim 28 \times \sim 28$  km), ensuring input of at least 20 datapoints and capturing the overall spatial distribution of *P. ramorum* records. Based on this stratification, 25 points were used to build the models and 43 points for testing predictive accuracy, with models trained and results mapped at the finer  $0.08^\circ$  resolution.

Raster coverages (again, pixel size  $0.08^\circ$ ) were assembled summarizing geographic distributions for 35 species of red oaks north of Mexico, as follows: *Quercus acerifolia*, *Q. agrifolia*, *Q. arkansana*, *Q. buckleyi*, *Q. coccinea*, *Q. ellipsoidalis*, *Q. emoryi*, *Q. falcata*, *Q. georgiana*, *Q. graciliformis*, *Q. gravesii*, *Q. hemisphaerica*, *Q. hypoleucoides*, *Q. ilicifolia*, *Q. imbricaria*, *Q. incana*, *Q. inopina*, *Q. kelloggii*, *Q. laevis*, *Q. laurifolia*, *Q. marilandica*, *Q. myrtifolia*, *Q. nigra*, *Q. pagoda*, *Q. palustris*, *Q. parvula* var. *shrevii*, *Q. phellos*, *Q. pumila*, *Q. robusta*, *Q. rubra*, *Q. shumardii*, *Q. texana*, *Q. velutina*, *Q. viminea* (Nixon,



**Figure 1** Predicted potential distribution for *Phytophthora ramorum* in California counties, based on known occurrence points (shown as yellow squares, inset) and three distinct environmental data sets, generated using the Genetic Algorithm for Rule-set Prediction (model outputs, which ranged 0–10, were simply added to produce this visualization). Models indicate the potential for spread more broadly along the coast, as well as to the western slopes of the Sierra Nevada. Shading denotes the number of models predicting potential presence in a particular pixel.

1980) and *Q. tardifolia*, but excluding *Q. wislizeni*, as it has not yet been reported as a foliar or bark canker host of *P. ramorum* (Dodd *et al.*, 2004). Two non-*Quercus* host trees, *L. densiflorus* and *Arbutus menziesii*, were also included. Distributional maps of 24 oaks and the two non-oak hosts were available as shapefiles (<http://climchange.cr.usgs.gov/data/atlas/little/>), which were converted into raster format. For the 11 remaining oak species, comparable datasets were created based on published range maps and distributional information (<http://hua.huh.harvard.edu/FNA/>; <http://www.csd1.tamu.edu/FLOTA/tracy/main1.html>; <http://www.calflora.org>). All host distributions were overlaid and aggregated over North America to derive a coverage of potential host tree distribution and used as a landscape mask to define the spatial extent across which SOD might be expected to occur in relation to host tolerances (see below).

Topographic and climatic inputs were from the Hydro-1 K dataset (U.S. Geological Survey; USGS) data describing topography (aspect, elevation, flow accumulation, flow direction, topographic index; <http://edcdaac.usgs.gov/gtopo30/hydro/>), monthly normalized difference vegetation index (NDVI) from years 1996–2000 derived from advanced very high resolution radiometer (AVHRR) imagery (<http://edc.usgs.gov/geodata/>), and yearly mean values (1961–1990) of diurnal temperature range, ground frost frequency, minimum, maximum and mean temperatures, precipitation, solar radiation, vapour pressure, and wet-day frequency drawn from the Intergovernmental Panel on Climate Change (IPCC) data

warehouse (<http://ipcc-ddc.cru.uea.ac.uk/>). USGS and NDVI data were  $0.01 \times 0.01^\circ$  base resolution ( $1 \times 1$  km), generalized to  $0.08 \times 0.08^\circ$  ( $8 \times 8$  km) for analysis; IPCC data were  $0.5 \times 0.5^\circ$  resolution ( $50 \times 50$  km) subdivided to  $0.08 \times 0.08^\circ$  ( $8 \times 8$  km) for analysis. Each of the three environmental data sources provides a distinct sort of information: topography summarizes landscape form; climatic data inform about the atmospheric conditions above that landscape; and NDVI provides a view of seasonal patterns of greenness across landscapes (obviously at least in part a function of topography and climate). NDVI can be used to infer climatic conditions, but tends to correlate well only with minimum temperature, mean temperature and precipitation, and soil moisture (Wang *et al.*, 2001; Adegoke & Carleton, 2002).

Based on these tradeoffs, five distinct ecological niche models were generated, differing in the type and spatial extent of environmental variables: NDVI + topography (constrained within the spatial distribution of potential host trees), NDVI + topography, climate + topography modelled over all of North America, and regional models of the latter two combinations. Regional models were based on environments represented within a 500-km buffer around known occurrence points.

For each set of models, the 10 best replicate GARP runs were chosen based on the combination of omission and commission error components calculated from the independent testing data mentioned above (Anderson *et al.*, 2003). These 'best subsets' of models were summed to produce maps of potential SOD distribution. In these

**Table 1** Receiver operator characteristic area under curve (ROC-AUC) scores evaluating predictive ability of ecological niche models presented herein, including the statistical significance (*P*) and standard errors (SE) for the five *Phytophthora ramorum* ecological niche model sets. Predictivity of all models was significantly better than random ( $P < 0.05$ ), but the climate + topography models were less able to discriminate between presences and absences in the independent test dataset

Model input	ROC-AUC	<i>P</i>	SE
NDVI + topography	0.824	$< 0.01E^{-15}$	0.038
NDVI + topography (500-km buffer)	0.831	$< 0.01E^{-15}$	0.042
NDVI + topography (landscape mask)	0.830	$< 0.01E^{-15}$	0.041
Climate + topography	0.672	$3.37E^{-4}$	0.048
Climate + topography (500-km buffer)	0.725	$3.92E^{-6}$	0.051

summary outputs, values ranged from 0 (areas where no models predicted potential presence) to 10 (areas where all models predicted potential presence). In general, areas in which all best-subsets models agreed in predicting potential presence were focussed on. The best-subset models were then projected to other regions to outline potential geographic distributions for the pathogen in the USA and worldwide (Peterson & Vieglais, 2001; Peterson, 2003). This approach has seen extensive testing in a variety of invasive species systems, and has shown excellent abilities in predicting the geographic course of already-established invasions (Peterson & Vieglais, 2001; Papes & Peterson, 2003; Peterson, 2003; Peterson *et al.*, 2003a,b).

To evaluate the accuracy of the GARP predictions, the area under the receiver operating characteristic (ROC) curve of the model's predictive performance was determined. An ROC curve represents the relationship between the probabilities of a model correctly classifying true positives (sensitivity; Y-axis) and incorrectly classifying false positives ( $1 - \text{specificity}$ ; X-axis) (Fielding & Bell, 1997). If these probabilities are equal, the model has no discriminatory power and will perform no better than random; this relationship is represented as a 'line of no information', a 1:1 relationship between X and Y axes (slope = 1) under which the area is 0.5 (Hanley & McNeil, 1982). As such, the area under curve (AUC) is the probability at which the model will make a successful classification, ranging from random (0.5) to perfect discrimination (1.0). AUC values of 0.81–1.00 are considered robust (Swets, 1988).

GARP AUC was calculated using a Wilcoxon test (Hanley & McNeil, 1982) to compare the number of landscape pixels where 0, 1, 2, etc., best-subset models predicted *P. ramorum* potential presence vs. the number of those pixels where the phytopathogen actually occurred (testing data;  $n = 25$ ). The statistical significance of each model set was determined with a Z-test of the GARP AUC versus that of a random model (0.5; Hanley & McNeil, 1982).

## Results and discussion

In general, the results of GARP analyses of the different environmental datasets yielded similar geographic

predictions. ROC analysis showed that models based on NDVI values performed particularly well, in contrast to the models based on just climate + topography (Table 1). The close similarity of results based on very different environmental datasets (Fig. 1) suggests that the predictions are at least robust to choice of environmental datasets. In general, these predictions indicate ample potential for spread north and south along the coast, as well as potential for spread to the western slopes of the Sierra Nevada.

To date, three different approaches have been used to model SOD distribution in California. Meentemeyer *et al.* (2004) applied a rule-based model, and predicted highest risk of SOD establishment in the state's northern and central coastal range between Del Norte and Santa Cruz counties (see Fig. 1 for county locations), and moderate risk in the foothills of the northern Sierra Nevada and in scattered patches of the southern coastal range from Santa Barbara county south to San Diego county. Support vector machines (svm) modelling by Guo *et al.* (2005) predicted areas of SOD risk in the central coastal range, from Mendocino county to Santa Barbara county, and the northern and central foothills of the Sierra Nevada. Venette & Cohen (2006) used a climate-matching model (CLIMEX) based on observed *P. ramorum* environmental tolerances, predicting favorable climatic conditions between Del Norte and Santa Cruz counties, and in the Sierra Nevada from Shasta county south to Tuolumne county. The GARP model results are most similar to the svm outputs, but differ from all three approaches by predicting greater SOD potential distribution across more of the Sierra Nevada foothills and in San Diego county (Fig. 1).

Projecting the models across North America indicated highest potential for establishment in the southeastern USA (Fig. 2a). Previous efforts over the contiguous USA were made using CLIMEX (Venette & Cohen, 2006) and a set of five different approaches: rule-based, logistic regression, classification and regression trees, GARP and svm (Kelly *et al.*, in press). The CLIMEX, rule-based, and logistic regression models showed a broad SOD potential distribution across the eastern United States, with highest risk in the southeastern states, whereas the Kelly *et al.* (in press) GARP and CART models demonstrated a band of highest risk spanning the middle southeast from the central Great Plains to the mid-Atlantic coast. The GARP outputs described here most closely matched the region predicted by the svm model (Kelly *et al.*, in press), which extends from the mid-Atlantic region southeast to eastern Texas.

Projections onto Europe indicated the northern Iberian Peninsula, the south and west of France, and central Italy as areas of highest model overlap (Fig. 3a). Model predictions coincided well with observed *P. ramorum* distributional areas in UK forests (primarily Wales and Cornwall; Department for Environment, Food & Rural Affairs, 2006). For the UK, CLIMEX and rule-based models of *P. ramorum* potential distribution showed high potential risk across the west and south, with hotspots in southern Wales and Cornwall (<http://www.defra.gov.uk/planh/>

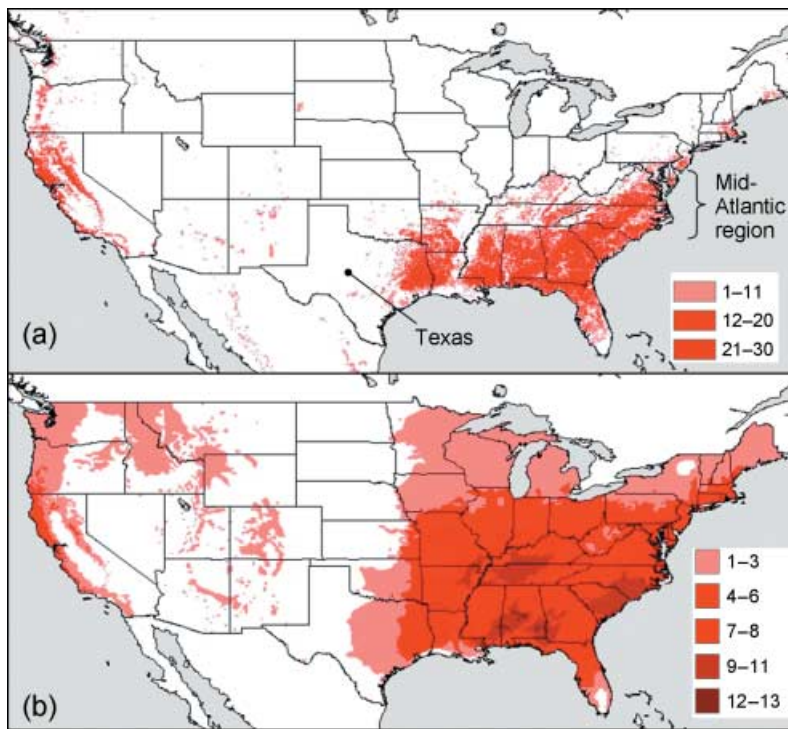


Figure 2 (a) Potential distribution of *Phytophthora ramorum* in North America, based on known California occurrences and three ecological datasets (as in Figure 1). Shading denotes the number of models predicting potential presence in a landscape. (b) Red oak (*Quercus* Section *Lobatae*) species richness in USA.

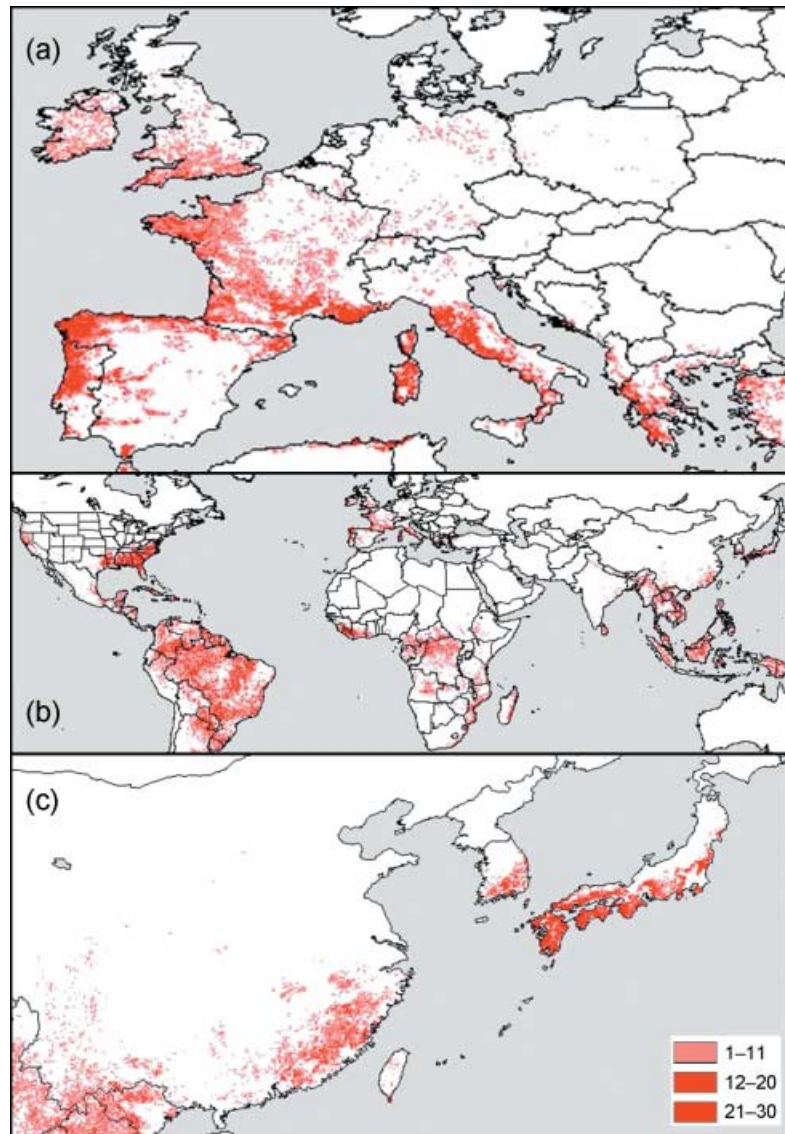
ramorum/meeting/webbinm.pdf; <http://www.defra.gov.uk/plant/ram2005/webber.pdf>).

Disagreement among the various modelling approaches may be the result of differences in computational approach, predictor variables, spatial/temporal scale, or combinations thereof. Discrepancies among models highlight uncertainties in predicting *P. ramorum* distribution (e.g. extent of potential range south of California's central coast). Conversely, consensus among modelling efforts strengthens the case for the validity of a prediction, and is a key consideration in evaluating an ENM's predictive ability when projected onto another region (i.e. beyond California). In this sense, for instance, the repeated conclusion that the southeastern USA is particularly vulnerable (Venette & Cohen, 2006; Kelly *et al.*, in press; this study) should be a clear indication of high risk of establishment there. Mexico (Fig. 2a,b), which is generally considered a key centre of endemism for oaks (Nixon, 1993), should also be regarded as at considerable risk. Foliar hosts mediate the transmission of *P. ramorum* (Garbelotto *et al.*, 2003), and the southeastern USA has high species richness among genera of known hosts (e.g. *Rhododendron*, *Vaccinium*) and potential hosts (e.g. Ericaceae: *Kalmia*, *Gaylussacia*) (<http://plants.usda.gov/>). This coincidence of ecological potential (e.g. in terms of climate) and presence of appropriate hosts indicates great potential for invasiveness in these regions.

Invasive plant pathogens have caused dramatic changes to native North American ecosystems, with substantial ecological and socioeconomic impacts (Crooks, 2002; Gilbert, 2002). Chestnut blight (*Cryphonectria parasitica*) and Dutch elm disease (*Ophiostoma* spp.) essentially eliminated the American chestnut (*Castanea dentata*) and

American elm (*Ulmus americana*), respectively, from eastern deciduous forests (Keever, 1953; Karnosky, 1979). With the loss of chestnut and elm, red oaks became dominant, but the reduced species diversity and habitat heterogeneity of these stands provide fewer barriers to transmission of oak phytopathogens (Wilson, 2001). Of the 35 red oak species in the USA, 18 occur primarily in the southeast (Fig. 2b; *Q. acerifolia*, *Q. arkansana*, *Q. buckleyi*, *Q. falcata*, *Q. georgiana*, *Q. hemisphaerica*, *Q. imbricaria*, *Q. incana*, *Q. inopina*, *Q. laevis*, *Q. laurifolia*, *Q. marilandica*, *Q. myrtifolia*, *Q. nigra*, *Q. phellos*, *Q. pumila*, *Q. shumardii* and *Q. texana*). Many of these southeastern oaks are regional endemics or microendemics with relatively limited distributions (Jensen, 1997) that may predispose them to extinction if exposed to a novel pathogen (e.g. Terborgh, 1974). Oak seed mast is an important food for wildlife, the reduction of which could have substantial effects on the physical and trophic structure of eastern forests (McShea & Healy, 2002; Rizzo & Garbelotto, 2003). Red oaks are the most important commercial hardwood species in the southeastern USA (Howard, 2001; Johnson, 2001), and the regional timber industry may suffer significant economic impacts if *P. ramorum* becomes established. The southeastern USA appears to be at greatest risk of *P. ramorum* establishment and spread, given the region's appropriate ecological conditions, high richness of potential host plants, and relatively limited heterogeneity of deciduous forests.

The assessment made in this study is based on the observed geographic distribution of North American *P. ramorum* populations (the A2 mating type). To date, all *P. ramorum* isolates from North American forests have



**Figure 3** Potential distribution of *Phytophthora ramorum* in Europe (a), worldwide (b), and in eastern Asia (c). Predictions are based on known Californian occurrences and three ecological datasets, and were generated using the Genetic Algorithm for Rule-set Prediction (as in Figure 1). Shading denotes the number of models predicting potential presence in a landscape.

been A2, but the A1 type (representative of European populations) has been detected in US and Canadian nurseries (Ivors *et al.*, 2004; Cave *et al.*, 2005; Rizzo *et al.*, 2005). The A1 and A2 mating types have different genotypic and phenotypic characteristics (Ivors *et al.*, 2004), and sexual recombination between North American and European *P. ramorum* lineages could alter their ecology (e.g. environmental tolerances, range of suitable host species, and/or level of aggressiveness; Ivors *et al.*, 2004; Cave *et al.*, 2005; Rizzo *et al.*, 2005).

Determining the geographic source of *P. ramorum* and sampling 'native' genomes would greatly enrich efforts towards understanding its genetic diversity, ecological amplitude, and the mechanisms of disease resistance in native-range hosts (Rizzo & Garbelotto, 2003; Ivors *et al.*, 2004; Cave *et al.*, 2005; Rizzo *et al.*, 2005). Globally, regions of highest model convergence were widely scattered (Fig. 3b), but the most geographically coherent

and consistently predicted potential source of *P. ramorum* was eastern Asia (Fig. 3c), a region speculated as the geographic origin of the pathogen (Brasier *et al.*, 2004, 2005). Southern Japan (Kyushu, Shikoku, southern Chugoku, Kansai, southern Chubu and northwest Kanto prefectures), southern South Korea (Kyongsang-nampo province), and eastern China (Fujian and Yunnan provinces) were identified as high-priority areas for the search for native *P. ramorum* populations. These regions are within the ecological niche of the invasive Californian populations, have a diversity of known *P. ramorum* hosts (e.g. Ericaceae in Yunnan province; Brasier *et al.*, 2004), and are well known as a source of species invasive to North America (National Research Council, 2002).

Finally, and most generally, *P. ramorum* distribution and ecology was modelled at least partly in the absence of information regarding host distributions and the importance of biotic interactions in shaping the species'

distribution. This approach stresses a key unresolved question in the broader field of community ecology (Ricklefs & Schluter, 1993) – what are the roles of interactions in shaping species' distributions as a function of spatial scale and resolution? Although the suggestion has been made that autecology will determine coarse-scale distributions of species, whereas interactions will affect fine-scale distributions (Soberón & Peterson, 2005), much additional research is needed to resolve the question. In the present case, final and definitive testing will require considerable time, as *P. ramorum* spreads across regions and demonstrates its full distributional potential.

### Acknowledgements

The views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the US Environmental Protection Agency. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

### References

- Adegoke JO, Carleton AM, 2002. Relations between soil moisture and satellite vegetation indices in the U.S. corn belt. *Journal of Hydrometeorology* 3, 395–405.
- Anderson RP, Lew D, Peterson AT, 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling* 162, 211–32.
- Brasier CM, Beales PA, Kirk SA, Denman S, Rose J, 2005. *Phytophthora kernoviae* sp. nov., an invasive pathogen causing bleeding stem lesions on forest trees and foliar necrosis of ornamentals in the UK. *Mycological Research* 109, 853–9.
- Brasier CM, Denman S, Brown A, Webber JA, 2004. Sudden oak death (*Phytophthora ramorum*) discovered on trees in Europe. *Mycological Research* 108, 1107–10.
- Cave GL, Randall-Schadel B, Redlin SC, 2005. *Risk Analysis for Phytophthora ramorum Werres, de Cock & In't Veld, Causal Agent of Phytophthora Canker (Sudden Oak Death), Ramorum Leaf Blight, and Ramorum Dieback*. Raleigh, NC, USA: U.S. Department of Agriculture, [http://www.aphis.usda.gov/ppq/ispmp/ramorum/pramorumpra05-05-05.pdf].
- Crooks JA, 2002. Characterizing ecosystem-level consequences of biological invasions: the role of ecosystem engineers. *Oikos* 97, 153–66.
- Department for Environment, Food & Rural Affairs, 2006. Findings of *Phytophthora ramorum* on plants growing in established gardens, woods and other wild sites in the UK 2002 to 2006. [http://www.defra.gov.uk/plant/graphics/sod.pdf].
- Dodd RS, Hüberli D, Douhovnikoff V, Harnik TY, Afzal-Rafii Z, Garbelotto M, 2004. Is variation in susceptibility to *Phytophthora ramorum* correlated with population genetic structure in coast live oak (*Quercus agrifolia*)? *New Phytologist* 165, 203–14.
- Elith J, Graham CH, Anderson RP *et al.*, 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29, 129–51.
- Erwin DC, Ribeiro OK, 1996. *Phytophthora Diseases Worldwide*. St. Paul, MN, USA: APS Press.
- Fielding AH, Bell JF, 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24, 38–49.
- Fry WE, Goodwin SB, 1997. Re-emergence of potato and tomato late blight in the United States. *Plant Disease* 81, 1349–57.
- Garbelotto M, Davidson JM, Ivors K *et al.*, 2003. Non-oak native plants are main hosts for sudden oak death pathogen in California. *California Agriculture* 57, 18–23.
- Gilbert GS, 2002. Evolutionary ecology of plant diseases in natural systems. *Annual Review of Phytopathology* 40, 13–43.
- Grinnell J, 1917. The niche-relationships of the California thrasher. *Auk* 34, 427–33.
- Guisan A, Theurillat JP, 2000. Equilibrium modeling of alpine plant distribution and climate change: how far can we go? *Phytocoenologia* 30, 353–84.
- Guo Q, Kelly M, Graham C, 2005. Support vector machines for predicting distribution of sudden oak death in California. *Ecological Modelling* 182, 75–90.
- Hanley JA, McNeil BJ, 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143, 29–36.
- Howard JL, 2001. *U.S. Timber Production, Trade Consumption, and Price Statistics 1965 to 1999*. Madison, WI, USA: U.S. Department of Agriculture: Research Paper FPL-RP-595.
- Ivors KL, Hayden KJ, Bonants PJM, Rizzo DM, Garbelotto M, 2004. AFLP and phylogenetic analyses of North American and European populations of *Phytophthora ramorum*. *Mycological Research* 108, 378–92.
- James FC, McCulloch CE, 2002. Predicting species presence and abundance. In: Scott JM, Heglund PJ, Morrison ML *et al.*, *Predicting Species Occurrences: Issues of Accuracy and Scale*. Washington DC, USA: Island Press, 461–6.
- Jensen RJ, 1997. *Quercus* Linnaeus sect. *Lobatae* Loudon. In: Flora of North America Editorial Board, ed. *Flora of North America*. Oxford, UK: Oxford University Press, 447–68.
- Johnson TG, 2001. *United States Timber Industry – an Assessment of Timber Product Output and Use, 1996*. Asheville, NC, USA: U.S. Department of Agriculture: General Technical Report SRS-45.
- Jules ES, Kauffman MJ, Ritts WD, Carroll AL, 2002. Spread of an invasive pathogen over a variable landscape: a non-native root rot on Port Orford cedar. *Ecology* 83, 3167–81.
- Kadmon R, Farber O, Danin A, 2003. A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications* 13, 853–67.
- Karnosky DF, 1979. Dutch elm disease: a review of the history, environmental implications, control, and research needs. *Environmental Conservation* 6, 311–22.
- Keever C, 1953. Present composition of some stands of the former oak-chestnut forest in the southern Blue Ridge Mountains. *Ecology* 34, 44–5.
- Kelly M, Shaari D, Guo Q, Liu D, in press. *Modeling Risk for SOD Nationwide: What Are the Effects of Model Choice on Risk Prediction?* Albany, CA, USA: U.S. Department of Agriculture: General Technical Report PSW-GTR-xxx. [http://nature.berkeley.edu/commtf/html/modeling\_phytophthora\_ramorum.html]

- Lek S, Delacoste M, Baran P, Dimopoulos I, Lauga J, Aulagnier S, 1996. Application of neural networks to modelling nonlinear relationships in ecology. *Ecological Modelling* **90**, 39–52.
- McShea WJ, Healy WM, 2002. *Oak Forest Ecosystems*. Baltimore, MD, USA: Johns Hopkins University Press.
- Meentemeyer R, Rizzo D, Mark W, Lotz E, 2004. Mapping the risk of establishment and spread of sudden oak death in California. *Forest Ecology and Management* **200**, 195–214.
- National Research Council, 2002. *Predicting Invasions of Nonindigenous Plants and Plant Pests*. Washington, DC, USA: National Academy Press.
- Nixon KC, 1980. *A Systematic Study of Quercus parvula Greene on Santa Cruz Island and Mainland California*. Santa Barbara, CA, USA: University of California, MS thesis.
- Nixon KC, 1993. The genus *Quercus* in Mexico. In: Bye TP, Lot RA, Fa JE, eds. *Biological Diversity of Mexico: Origins and Distribution*. New York, USA: Oxford University Press, 447–58.
- Olden JD, 2000. An artificial neural network approach for studying phytoplankton succession. *Hydrobiologia* **436**, 131–43.
- Papes M, Peterson AT, 2003. Predictions of the geographic course of the invasion of *Ageratina adenophora* in China. *Journal of Wuhan Botanical Research* **21**, 137–42.
- Peterson AT, 2003. Predicting the geography of species' invasions via ecological niche modeling. *Quarterly Review of Biology* **78**, 419–34.
- Peterson AT, Ball LG, Cohoon KP, 2002a. Predicting distributions of Mexican birds using ecological niche modelling methods. *Ibis* **144**, E27–E32.
- Peterson AT, Cohoon KC, 1999. Sensitivity of distributional prediction algorithms to geographic data completeness. *Ecological Modelling* **117**, 159–64.
- Peterson AT, Papes M, Kluza DA, 2003a. Predicting the potential invasive distributions of four alien plant species in North America. *Weed Science* **51**, 863–8.
- Peterson AT, Scachetti-Pereira R, Kluza DA, 2003b. Assessment of invasive potential of *Homalodisca coagulata* in western North America and South America. *Biota Neotropica* **3**, [<http://www.biotaneotropica.org.br/v3n1/en/abstract?article+BN0070301200>].
- Peterson AT, Stockwell DRB, Kluza DA, 2002b. Distributional prediction based on ecological niche modeling of primary occurrence data. In: Scott JM, Heglund PJ, Morrison ML, eds. *Predicting Species Occurrences: Issues of Scale and Accuracy*. Washington, DC, USA: Island Press, 617–23.
- Peterson AT, Vieglais DA, 2001. Predicting species invasions using ecological niche modeling. *Bioscience* **51**, 363–71.
- Ricklefs RE, Schluter D, 1993. *Species Diversity in Ecological Communities: Historical and Geographical Perspectives*. Chicago, USA: University of Chicago Press.
- Rizzo DM, Garbelotto M, 2003. Sudden oak death: endangering California and Oregon forest ecosystems. *Frontiers in Ecology and the Environment* **1**, 197–204.
- Rizzo DM, Garbelotto M, Davidson JM, Slaughter GW, Koike ST, 2002. *Phytophthora ramorum* as the cause of extensive mortality of *Quercus* spp. and *Lithocarpus densiflorus* in California. *Plant Disease* **86**, 205–14.
- Rizzo DM, Garbelotto M, Hansen E, 2005. *Phytophthora ramorum*: integrative research and management of an emerging pathogen in California and Oregon forests. *Annual Review of Phytopathology* **43**, 309–335.
- Rushton SP, Ormerod SJ, Kerby G, 2004. New paradigms for modelling species distributions? *Journal of Applied Ecology* **41**, 193–200.
- Scott JM, Heglund PJ, Morrison ML *et al.*, 2002. *Predicting Species Occurrences: Issues of Accuracy and Scale*. Washington DC, USA: Island Press.
- Soberón J, Peterson AT, 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics* **2**, 1–10.
- Stockwell DRB, 1999. Genetic algorithms II. In: Fielding AH, ed. *Machine Learning Methods for Ecological Applications*. Boston, MA, USA: Kluwer Academic Publishers, 123–44.
- Stockwell DRB, Noble IR, 1992. Induction of sets of rules from animal distribution data: a robust and informative method of analysis. *Mathematics and Computers in Simulation* **33**, 385–90.
- Stockwell DRB, Peters D, 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science* **13**, 143–58.
- Stockwell DRB, Peterson AT, 2002. Effects of sample size on accuracy of species distribution models. *Ecological Modelling* **148**, 1–13.
- Swets JA, 1988. Measuring the accuracy of diagnostic systems. *Science* **240**, 1285–93.
- Swihart RK, Gehring TM, Kolozsvary MB, Nupp TE, 2003. Responses of 'resistant' vertebrates to habitat loss and fragmentation: the importance of niche breadth and range boundaries. *Diversity and Distributions* **9**, 1–18.
- Terborgh J, 1974. Preservation of natural diversity: the problem of extinction prone species. *BioScience* **24**, 715–22.
- Venette RC, Cohen SD, 2006. Potential climatic suitability for establishment of *Phytophthora ramorum* within the contiguous United States. *Forest Ecology and Management* **231**, 18–26.
- Wang J, Price KP, Rich PM, 2001. Spatial patterns of NDVI in response to precipitation and temperature in the central Great Plains. *International Journal of Remote Sensing* **22**, 3827–44.
- Werres S, Marwitz R, Man in't Veld WA *et al.*, 2001. *Phytophthora ramorum* sp. nov: a new pathogen on *Rhododendron* and *Viburnum*. *Mycological Research* **10**, 1155–65.
- Weste G, 2003. The dieback cycle in Victorian forests: a 30-year study of changes caused by *Phytophthora cinnamomi* in Victorian open forests, woodlands and heathlands. *Australasian Plant Pathology* **32**, 247–56.
- Wilson AD, 2001. Oak wilt: a potential threat to southern and western oak forests. *Journal of Forestry* **99**, 4–11.