Received: 11 November 2022 | Revised: 17 March 2023 | Accepted: 23 March 2023

DOI: 10.1002/ecs2.4573

ECOSPHERE
AN ESA OPEN ACCESS JOURNAL

ARTICLE
Disease Ecology

# Multivariate Bayesian analysis to predict invasiveness of *Phytophthora* pathogens

**Bruce G. Marcot**[1] | **Peter Scott**[2,3] | **Treena I. Burgess**[3]

[1]USDA Forest Service, Pacific Northwest Research Station, Portland, Oregon, USA

[2]Department of Primary Industries and Regional Development, Sustainability and Biosecurity, Perth, Western Australia, Australia

[3]Harry Butler Institute, Murdoch University, Perth, Western Australia, Australia

**Correspondence**
Bruce G. Marcot
Email: bruce.marcot@usda.gov

**Funding information**
U.S. Forest Service, Pacific Northwest Research Station; Department of Primary Industries and Regional Development, Perth, Western Australia

**Handling Editor:** Charles E. Mitchell

**Abstract**

Global concerns are many for the invasive impacts of *Phytophthora* pathogens on native vegetation, agriculture, nurseries, and urban parks and gardens. We compiled a database of 32 traits on 204 species of *Phytophthora* including data on each species' taxonomy (clade and subclade), historical knowledge (years since first described), impacted ecosystems, microenvironments inhabited, dispersal mode, physiology, and morphology. Drawing from approximately 11,394 unique host, pathogen, and country plant disease records from GenBank and other sources, we calculated potential invasiveness of 103 better studied species from cluster relationships. We used the species data to create a Bayesian network model predicting the degree and probability of invasiveness of individual *Phytophthora* species. Model calibration testing resulted in <1% error rate in classifying invasiveness categories of well-known species. We applied the model to predict the potential invasiveness of 101 other species with unknown invasiveness dynamics. The model can also be used to predict the invasive risk of other poorly studied and newly identified *Phytophthora* species, and the general modeling approach can be used for other pests and pathogens, to advise land and resource managers to thwart potential invasions before they occur or intensify.

**KEYWORDS**
agriculture, Bayesian network, forests, invasive, natural ecosystems, *Phytophthora*

## INTRODUCTION

The global community is concerned about the continued and new introductions and impacts of *Phytophthora* pathogens responsible for numerous serious plant diseases in native vegetation. Examples include the pathogen *P. ramorum* causing sudden oak death in coast live oak (*Quercus agrifolia*) and tanoak (*Notholithocarpus densiflorus*) in western United States (Meentemeyer et al., 2011; Rizzo & Garbelotto, 2003); *Phytophthora* spp. causing dieback of Proteaceae in Australia and South Africa (Cahill et al., 2008; Nagel et al., 2013); *P. agathidicida* causing dieback of ancient kauri trees (*Agathis australis*) of great cultural value in northern New Zealand (Bradshaw et al., 2020); and many more (e.g., Fernández-Habas et al., 2019; Ruffner et al., 2019). These pathogens are irreversibly changing the ecosystems they impact. *Phytophthora* pathogens within plant production systems also cause considerable economic losses, and the cost of managing these diseases is integrated into

these production systems (Panth et al., 2020). Some of the most severe economic impacts of *Phytophthora* diseases are late blight of potato (Cooke et al., 2011), black pod disease of cacao (Acebo-Guerrero et al., 2012), root rot in avocado (Ramírez-Gil et al., 2017), root rot in soybean (Dorrance, 2018), root and basal rot of capsicum (Granke et al., 2012), and root, collar, and fruit rot of citrus (Graham & Feichtenberger, 2015).

In their natural environment, *Phytophthora* species provide a range of functions including supporting nutrient decomposition and coexistence of plant species (Hansen, 2015; Laliberté et al., 2015; Marano et al., 2015). Many *Phytophthora* species can be characterized as synanthropic, where individual species may be described from a particular location or region but can disperse more widely because of associations with human activities such as being carried with plants or soil attached to machinery or on boot treads into otherwise virgin environments.

Of central concern is determining their degree of potential harm to native vegetation and plant production systems, prompting surveys (e.g., Burgess, White, et al., 2017; Parke et al., 2020), which, unfortunately, are generally instituted only after a *Phytophthora* disease becomes established and evident. Land managers need a means of predicting the potential degree of invasiveness in new environments and locations before introductions and impacts. Hence, the current study addresses this need to determine the biosecurity risk of *Phytophthora* species based on their ecology, taxonomy, morphology, physiology, and traits associated with their known host range distribution and ecosystem occurrence. We aimed to develop and provide a tool that biologists and land managers could use to determine the potential risk of invasiveness of *Phytophthora* species in new locations.

Analyses of invasiveness and biosecurity risk of *Phytophthora* species have mainly focused on estimating the potential range of species that are known to have high impacts (Burgess, Scott, et al., 2017; Ireland et al., 2013; Schoebel et al., 2014; Tsykun et al., 2022). However, many new significant diseases are caused by *Phytophthora* species that are either new to science or have a negligible impact until they are found to be causing major damage.

Gaydos et al. (2019) assessed the value of participatory epidemiological modeling of potential pathogen effects and control strategies for treating sudden oak death. Several studies have identified specific traits determining species distributions, such as the effects of topography on dispersal of *P. cinnamomi* (Cardillo et al., 2018), density of host plants, tree diversity, and pathogen load on the occurrence of *P. ramorum* (Cobb et al., 2012; Dillon & Meentemeyer, 2019), dispersal of *P. ramorum* by rainwater and hiker's boots (Davidson et al., 2005), and influence of climate on blight from *P. infestans* (Garrett et al., 2009).

Fewer studies still have addressed global distributions. Barwell et al. (2021) used *Phytophthora* species attributes such as host range, habitat physiological and morphological traits of the individual species, and phylogenetic generalized linear mixed models to predict the number of countries reached, latitudinal limits, and host range. By comparing known to predicted values from the models, several newly described species were predicted to have higher impacts than currently observed. The analysis of Scott et al. (2019) is most relevant to the current work; they compiled a global database of *Phytophthora* occurrences, estimated global species richness, and predicted invasion and infection potential based on species biogeography and host range. The limitation of their study was that many countries, especially in the developing world, have few records and thus lower than expected species numbers. Thus, while they used a database to create a tool for comparing species present in two countries, the absence of a *Phytophthora* species in one country may reflect a lack of reporting rather than a genuine absence.

The more than 200 species of *Phytophthora* included in this study vary considerably in attributes and environments occupied. Many species are not well known, imparting greater uncertainty about their performance in novel environments and locations. Given this high variability and uncertainty, the modeling framework we chose had to be capable of using the best information available on better known species for predicting and testing their invasiveness, applying that foundation to predictions of poorer known species, and providing a basis for improving model predictions with new data. Thus, we selected an approach based on Bayesian statistics, specifically Bayesian networks (BNs), which uses the strength of evidence of prior information—in this case, data on the better known species—and explicitly denotes variability and uncertainty in the form of probabilities.

## METHODS

### *Phytophthora* disease records

Publicly available records of plant diseases involving *Phytophthora* species were collated to establish a global database of *Phytophthora* diseases. Reports of formally and provisionally described species for which host genus or species information was available were included until March 2022. Reports were sourced from plant disease databases and large collections including:

1. United States Department of Agriculture (USDA), global fungal database (Farr & Rossman, 2022) using the R package (rusda—Krah, 2016, $n = 8497$ reports),
2. Centre for Agriculture and Bioscience International (CABI),
3. The New Zealand Forest Research (Scion) forest health database (collected through correspondence = 1504 records),
4. Landcare Research, New Zealand fungi database (collected through correspondence = 1559 records),
5. Australia *Phytophthora* disease records (sourced from Burgess et al., 2021 = 7629 records), and
6. NCBI GenBank records (Supporting Data 1; see details below = 12,957 records).

Additional plant disease records not available in the above databases were manually downloaded (470 records). All plant disease records including published disease records, international databases, and GenBank accessions are included in Burgess and Scott (2023).

## Molecular records

All internal transcribed spacer (ITS) sequence data from *Phytophthora* were downloaded from GenBank (https://www.ncbi.nlm.nih.gov/, accessed on 22 March 2022) using the search terms *Phytophthora* [organism] AND internal transcribed spacer [any field]. Additionally, a blast search was conducted for a representative isolate of each described *Phytophthora* species to capture any additional sequences. This dataset consists of approximately 17,000 accessions. All data were uploaded into Geneious Prime 2019.2.3 (https://www.geneious.com/) and a data-cleaning process followed (for full explanation, see Burgess & Scott, 2023). Species identities were confirmed by comparison to sequences of type isolates available from IDPhy (https://idtools.org/id/phytophthora/) (Abad et al., 2023).

## Data cleaning

All records were checked for spelling errors and inconsistencies. Host plant taxonomic synonymies were resolved relevant to "The Plant List" (www.theplantlist.org) and "World Flora Online" (http://www.worldfloraonline.org/) and updated to current naming conventions using the R packages taxonstand (Cayuela et al., 2021) and WorldFlora (Kindt, 2021).

*Phytophthora* species taxonomic synonyms were checked against phylogenetic trees listed in the website *Phytophthora* ID tools (https://idtools.org/id/phytophthora/). Country and continent names were matched to current designations listed by the United Nations (https://www.un.org/en/about-us/member-states). The free software R was used for compiling and sorting disease records, statistical computations of invasiveness, and some graphics (R Core Team, 2022).

Records were included in the analysis only if the *Phytophthora* species, date, country, and host details (up to the genus level) were available. Where plant disease records had the same *Phytophthora* species, host records, country, and year on multiple data sources, only one disease record was used for the analysis. In total, 11,394 unique records on *Phytophthora* species, host, country, and year were used for the analysis, including the R code used for data cleaning (Burgess & Scott, 2023).

## Designation of confidence group

*Phytophthora* species were designated as belonging to either a "yes confident" (Y-confident) or a "no confident" (N-confident) group. The grouping indicates whether we have confidence in the disease and host reports associated with each species. The Y-confident group includes species for which the pathogen taxonomy has been clearly resolved when the disease reports were made and included species where molecular phylogeny was clearly defined and the species has been studied more intensely. The selection was primarily based on the year of description. Only five species described after 2013 (*P. fragariaefolia*, *P. nagaii*, *P. parvispora*, *P. chlamydospora*, and *P. versiformis*) were included in the Y-confident group. Most species described before 2010 were included; however, species that were commonly misnamed (occurring in a species complex or because of dubious issues with the nomenclature) were excluded (*P. pini*, *P. citricola*, *P. megasperma*, *P. alticola*, *P. parsiana*, and *P. sansomeana*).

The N-confident group included species where the pathogen taxonomy was not fully resolved when the disease reports were first made, or if the taxonomy was recently resolved and there are insufficient data to confirm the disease reports. Therefore, the N-confident group includes more newly described species and many species within recently resolved species complexes. The average number of years since species description was 41.2 years in the Y-confident group and 9.4 years in the N-confident group (Burgess & Scott, 2023). In total, 103 species were identified as belonging to the Y-confident group and 101 to the N-confident group.

## Calculating species invasiveness

An index of *Phytophthora* species invasiveness was calculated as described in the study by Scott et al. (2019).

Phylograms of the *Phytophthora* species were made using the Enhanced Heat Map: "Heatmaps2" function with the R package gplots (Warnes et al., 2022). The phylograms were made using normalized (center-scaled) indices including the:

1. number of *Phytophthora* reports by country (CP; country);
2. annual new species report rates by country (CRP; country id. rate);
3. number of *Phytophthora* reports (HP; host species); and
4. annual new species report rates (HRP; host id. rate).

The species branch height was determined using the "get_branches_heights" function of the dendextend package (Galili, 2015). The branch height for each species was normalized to range [0,1] using the below formula:

$$\text{Normalized score} = \frac{\text{Species branch height} - \min(\text{total branch height})}{\text{Max}(\text{total branch height}) - \min(\text{total branch height})}.$$

A normalized score of 1 denotes high probability of characteristics lending to pathogenic properties and invasiveness.

## *Phytophthora* traits

We collated trait data for 204 *Phytophthora* species, of which 195 are formally described, and 9 are provisionally named (Burgess & Scott, 2023). Phylogenetic placement into clade and subclade is widely accepted and was assigned as per Yang et al. (2017) and Bourret et al. (2018). The year refers to the number of years, as of 2022, since the species was first described, not the date a species was first isolated. The morphological and physiological variables were extracted from the tabular key available on IDPhy, where possible, while for newly described species the information was taken directly from the species description. If not provided in the manuscript, the oospore wall index was calculated from the average measurements of the oospore and the oospore wall as provided in the description or calculated based on images provided in the description. Where a temperature range was provided, the midpoint of the range was used. Traits can be linked to reproduction (sporangia, oospores, and chlamydospores), growth (mycelial growth rate and other cardinal temperature data), or survival (oospore wall thickness, hyphal swelling, chlamydospores, and temperature range for mycelial growth) (Table 1). Data for 170 species were published in the study by Barwell et al. (2021).

In addition, disease notes and reports were used to determine dispersal mode, disease traits (microenvironments infected), and impacted ecosystems. Disease traits were recovered from disease notes in the U.S. National Fungus Collections (Farr & Rossman, 2022) and from disease reports. These disease traits represent variation in the detectability of different types of disease symptoms. The disease traits also are ecologically important for invasion if they influence modes of dispersal (e.g., soilborne, airborne, waterborne) and transmission among hosts (Table 1), or if species that cause both root and foliar disease have more diverse mechanisms to overcome host defenses. Traits included may confer greater success at one or more stages during the invasion process of arrival (survival during transport), establishment (impacted ecosystems and disease caused), and spread (caducous).

## BN model development

We modeled invasiveness using BNs that allow inducing a model structure from empirical data and providing a basis for testing prediction accuracy and contributions of individual trait predictors. BNs are structured as a causal network of variables linked by conditional probabilities, whereby each variable is represented as a discrete set of states (categorical, ordinal, cardinal, or continuous value). Links (also called arcs) between variables (also called nodes) denote direct influences (first-order Markov processes). Probability distributions among states propagate through the network as uncertainties, eventually represented in the response variable(s) as a posterior probability distribution among the outcome states. BNs have been used widely in areas of natural resource modeling (McCann et al., 2006), biocontrol (Meurisse et al., 2021), predictions of presence of rare species including fungi (Marcot, 2006), and much more. An alternative approach is development of species distribution models (SDMs), such as the SDMs developed by Václavík and Meentemeyer (2009) of *P. ramorum* invasiveness in California. However, SDMs require extensive and spatially explicit data on presence, absence, and other species' habitat-selection attributes that would not be available on most *Phytophthora* species.

We developed our prediction model based on the suite of *Phytophthora* variables discussed above as covariates (predictors) and the invasiveness index as the response (prediction) variable (Table 1), using the Bayesian network modeling program Netica v6.09 (Norsys.com). The overall BN model development workflow was as follows: import and discretize the prediction variables (covariates) and the

**TABLE 1** Variables used in the Bayesian network model predicting *Phytophthora* species invasiveness.

| Variable group | Variable name and definition | States |
|---|---|---|
| Taxonomy | Phylogenetic clade | Clade codes |
| | Phylogenetic subclade | Subclade codes |
| Year | Years since first described: no. years as of 2022 since species taxon was first described | Continuous values |
| Measurable variables: physiology | Temperature minimum for mycelia growth: minimum temperature known for growth of mycelia (C) | Continuous values |
| | Temperature optimal for mycelia growth: optimal temperature known for growth of mycelia (C) | Continuous values |
| | Temperature maximum for mycelia growth: maximum temperature known for growth of mycelia (C) | Continuous values |
| | Temperature range for mycelia growth: overall range of temperatures for mycelia growth (C) | Continuous values |
| | Growth rate at optimal temperature: growth rate of a colony on an agar plate at optimal temperature (mm/day) | Continuous values |
| Measurable variables: morphology | Breeding system | HE = heterothallic, HO = homothallic, S = (functionally) sterile |
| | Papillate | P = papillate, SP = semi-papillate, NP = non-papillate |
| | Oospore wall | S, W, O = relates to the oogonia wall: smooth, wavy ornamented |
| | Oospore wall thickness | Continuous values |
| | Oospore wall index (ratio) | Continuous values |
| | Oospore base | TB = tapering base, NTB = no tapering base |
| | Chlamydospores | C = chlamydospore, NC = no chlamydospores |
| | Amphigynous, paragynous | A = amphigynous, P = paragynous, B = both |
| | Aplerotic plerotic | A = aplerotic, P = plerotic, B = both |
| | Caducous | C = caducous, NC = non-caducous |
| | Sporangiophore | UN = unbranched, SS = simple sympodial, CS = compound sympodial, DM = downy mildew |
| | Proliferation | P = proliferation, NP = no proliferation |
| | Hyphal swelling | H = hyphal swellings, NH = no hyphal swellings |
| Dispersal | Dispersal mode | Aerial, soil; soil; soil, water; water |
| Impacted ecosystems | Found in natural ecosystems | Found in natural ecosystems |
| | Found in urban parks, gardens, nurseries | Found in urban parks and gardens or nurseries |
| | Found in agriculture, perennial crops, forestry | Found in agriculture perennial crops or forestry |
| | Found in water | Found in water |
| | Found in agriculture annual crops | Found in agriculture annual crops |
| Microenvironments | Disease of leaves or fruit | Disease of leaves or fruit |
| | Stem canker | Disease of woody plants with cankers formed in the stem |
| | Disease or roots | Diseases of roots (fine or woody) |

*Note*: Source material includes IPPhy and species description and disease reports. Data for the morphological and physiological variables for 170 of the 204 were published previously (Barwell et al., 2021). There is also an updated tabular key on IDPhy (https://idtools.org/id/phytophthora/).

response variable into model nodes, each discretized with exclusive states; induce the causal network structure from the Y-confident species database ($n = 103$ spp.); parameterize the probability tables of all nodes from the Y-confident species database; test the classification (calibration) accuracy of the model against the Y-confident species database; conduct sensitivity analysis to determine the degree to which the response variable of invasiveness is sensitive to each covariate; using the sensitivity analysis results, determine which covariates may contribute little to model predictions and may be eliminated; eliminate low-contribution covariates and retest the classification accuracy of the model against the Y-confident species database; if the final classification accuracy (low error rate) is satisfactory, accept the final model structure and generalize the outer value ranges of the continuous variables; and then use the final model to predict invasiveness of the N-confident species set ($n = 101$ spp.), and explore conditions of selected species. Specific model development steps were as follows.

First, we imported the Y-confident species database into Netica, creating individual nodes for each variable. We retained all categorical values as discrete states and specified five levels of states for continuous-value variables. In BN models, for a given case sample size, a greater number of continuous-value states may increase precision but decrease the accuracy of representation and prediction; five levels were a fair balance given the number of species in the case Y-confident database ($n = 103$). We used the Y-confident case file database to discretize the continuous-value nodes, that is, to determine cutoff values to have each state represented by an approximately equal number of cases (species) with some uncertainty. This step included the invasiveness response variable, values of which were calculated for each Y-confident species as noted above.

Next, we inferred the causal network structure from the species database by implementing the machine learning Tree Augmented Naïve Bayes (TAN) algorithm (Jiang et al., 2012), which links the invasiveness response variable to the affector variables (covariates) as a predictive framework. This produces the so-called "naive Bayes" network structure where a specified output variable, here, invasiveness, is linked to all possible covariates. TAN also links covariates to one another if they are significantly correlated, which eliminates the need in this network structure for initially deleting correlated variables as is commonly done with frequentist multivariate modeling.

We then induced the values of all prior and conditional probability tables in the model by incorporating the Y-confident species database using the expectation maximization (EM) algorithm (Dlamini, 2011). In particular, appropriate selection of prior probabilities is an important step in Bayesian model development (Wesner & Pomeranz, 2021). EM is a convergent log-likelihood function that iteratively walks through the data cases and assigns conditional probability values in the model to best fit the relations of the response variable to its covariates. Where combinations of covariate state values were not present in the species database, Netica specifies uniform probability distributions denoting complete uncertainty for those situations, which may or may not occur in actuality. The result at this stage is a fully specified model.

Next, we tested the classification accuracy (error rate) of the model against the species database, using the Y-confident species set and its calculated invasiveness scores. This step amounts to model calibration testing, not validation against an independent dataset (NRC, 2012). Because there were too few species cases in the database to conduct cross-validation model development and testing (Marcot & Hanea, 2021), we opted for calibration classification accuracy testing. In the outcome invasiveness node, we used the state with the highest, dominant probability as the model prediction for each species. This testing provides a so-called confusion table denoting numbers of cases that were, and were not, correctly predicted as to their invasiveness score state. Well-calibrated models will have classification error rates <10% or so. We also calculated three other complementary measures of model calibration fit (Marcot, 2012): logarithmic loss, values of which range [0, infinity], with 0 denoting best model performance; quadratic loss, values of which range [0,2], also with 0 denoting best model performance; and spherical payoff, values of which range [0,1], with 1 denoting best model performance.

We then ran a sensitivity analysis on the model to determine the degree to which the invasiveness response variable was influenced by each covariate, calculating sensitivity as percent variance reduction in the response of the invasiveness score to each covariate (Marcot, 2012). We rank-ordered the results and identified which, if any, of the covariates contributed negligible sensitivity effects. We then removed those lowest sensitivity covariates, contributing <0.3% variance reduction, to simplify the model, including also removing the variable phylogenetic clade because it was fully redundant with variable phylogenetic subclade. Next, we restructured the model based on the remaining variables, reran the calibration classification accuracy analysis, and compared those error rates to those from the model with all variables included. If classification error remained low, we accepted the amended model as the final and more parsimonious structure.

We made two final edits to the revised model. First, to generalize the model, we expanded the lowermost and

uppermost values of all continuous-value nodes to encompass values that could be lower or higher than in the Y-confident case file dataset. This would not alter the model's calibration classification performance or sensitivity structure. Second, to make the model more understandable, we assigned word descriptions to the five states in the invasiveness outcome node: very low, low, moderate, high, and very high. Note that when the model is run on individual species, the expected value of the invasiveness score is still displayed at the bottom of that node.

We then ran the final, accepted model, structured and parameterized as above based on the Y-confident species set, to calculate invasiveness scores for the N-confident species. We present those results denoting species that may carry high potential for invasive dynamics. We then explored two species case examples to demonstrate species-specific considerations when using the results of our modeling tool.

## Next steps

We further discuss how the model can be used in real-world situations to predict the potential invasiveness of *Phytophthora* species detected in new locations and environments, and to identify which covariates may play key roles in each species' invasiveness score. We end with a discussion of how the model can be further refined and improved by incorporating new sightings and information on *Phytophthora* species, adding to the case files and updating the model probability parameters incrementally, and testing for updates to the model network structure, sensu the overall philosophy of Bayesian statistical analysis by accreting prior knowledge and adding new observations. The model and the databases used in this study are posted for general access and utilization.

# RESULTS

## *Phytophthora* country and host database

We assigned 14,820 GenBank accessions of ITS data to known *Phytophthora* species (Burgess & Scott, 2023). Country data were available for 12,254 accessions of known *Phytophthora* species, and of those, there were also host data for 6787 sequences (Burgess & Scott, 2023). We found that 16% of accessions submitted to GenBank were misnamed (Burgess & Scott, 2023). The most misnamed species were: *P. capsici* (many sequences align to *P. tropicalis*); *P. cryptogea* and *P. drechsleri* (many sequences align to *P. pseudocryptogea* or *P. kelmanii*); and *P. citricola* (many sequences align to *P. multivora*, *P. plurivora* or *P. pini*). When conducting GenBank searches, users must carefully check species identities based on the sequences of type isolates available from IDPhy (Abad et al., 2023). Data have been submitted to GenBank from 103 countries. The top 10 most common species are *P. nicotianae*, *P. cinnamomi*, *P. infestans*, *P. plurivora*, *P. palmivora*, *P. cactorum*, *P. capsici*, *P. lacustris*, *P. gonapodyides*, and *P. pseudocryptogea* (Burgess & Scott, 2023). The top 10 submitting countries are United States, Italy, India, China, Australia, Switzerland, Spain, Colombia, South Africa, and Poland (Burgess & Scott, 2023). In total, 11,394 unique *Phytophthora* species, host, country, and year plant disease records were used to determine the invasiveness of the *Phytophthora*, including 5876 unique disease records sourced from the GenBank records.

## Network model structure and calibration tests

The full, initial BN model constructed from the Y-confident species set included all 30 covariates and consisted of 59 links (Table 2). Based on the Y-confident species set, the values of the branch height invasiveness index in the outcome node states were partitioned as follows: very low 0 to <0.008; low 0.008 to <0.016; moderate 0.016 to <0.034; high 0.034 to <0.09; and very high 0.09 to 1. The model calculations of the 103 species in the Y-confident dataset resulted in 22 species rating very low invasiveness, 20 rating low, 20 rating moderate, 19 rating high, and 22 rating very high. This model produced a calibration classification error rate of about 11%, resulting from errors of predicted invasiveness score categories for 11 of the 103 species tested (Table 3).

Sensitivity analysis was then used to identify five covariates—oogonia features, found in roots, proliferation, oogonia base, and hyphal swelling—that contributed least (<0.3%) to variance reduction and thus least to

**TABLE 2** Attributes of the Bayesian network models predicting *Phytophthora* species invasiveness states.

| Network attribute | Initial full model | Final model |
|---|---|---|
| No. response variable nodes | 1 | 1 |
| No. covariate nodes | 30 | 25 |
| No. links between nodes | 59 | 47 |
| No. probability values in unconditional (prior) and conditional probability tables | 1380 | 2520 |

**TABLE 3** Results of calibration (classification accuracy) tests of the Bayesian network models predicting *Phytophthora* species invasiveness states.

| Calculated invasiveness state | Model-predicted invasiveness state | | | | |
|---|---|---|---|---|---|
| | Very low | Low | Moderate | High | Very high |
| Initial full model | | | | | |
| Very low | 19 | 1 | 1 | 0 | 1 |
| Low | 0 | 19 | 1 | 0 | 0 |
| Moderate | 1 | 0 | 18 | 1 | 0 |
| High | 0 | 0 | 1 | 17 | 1 |
| Very high | 2 | 0 | 0 | 1 | 19 |
| Final model | | | | | |
| Very low | 22 | 0 | 0 | 0 | 0 |
| Low | 0 | 20 | 0 | 0 | 0 |
| Moderate | 0 | 0 | 20 | 0 | 0 |
| High | 0 | 0 | 0 | 19 | 0 |
| Very high | 0 | 0 | 1 | 0 | 21 |

*Note*: Values are the number of species, listing classification accuracy results of the initial, full model with all 30 covariates, and of the final model with 25 covariates that omitted 5 covariates with low sensitivity scores or redundant contribution. Comparison of metrics of classification error rates and scoring rules.

predicting invasiveness scores given the contributions of all the other covariates (Figure 1). These five covariates, along with phylogenetic clade, were excised from the network, which was then restructured, reparameterized, and retested. Traits that contributed the most to model sensitivity (≥13% variance reduction) and to predicting invasiveness scores included phylogenetic subclade, breeding system, being found in urban parks, being found in water, and years since discovered (Figure 1; Appendix S1: Table S1).

The final, revised model (Figure 2) produced a calibration classification error rate of about 1%, with only one species (*P. chlamydospora*) misclassified (predicted invasiveness score of 0.544, state category is very high, compared with the calculated invasiveness score of 0.302, state category is moderate) (Table 4; Appendix S1: Table S2). The predictions for the 103 Y-confident species were nearly evenly distributed among the five invasiveness state outcome categories (Table 3), with 40 of the species (38%) ranking high or very high invasiveness.

Other model performance metrics also highlighted the improved performance of the modified network. Logarithmic loss and quadratic loss declined by 78% and 88%, respectively, and spherical payoff increased by 8%, with all these trends indicating much improved model performance (Table 3). The initial model with all 30 covariates included more nodes and links but a smaller number of probability table values than the trimmed final model with 24 covariates (Table 2). This pattern appropriately resulted from the final model, although having six fewer covariates, also being structured strictly from the empirical data with some nodes having a greater number of incoming links, which numerically increases the size of conditional probability tables (Marcot, 2012).

## Invasiveness predictions for N-confident species

Lastly, we ran the final model with the set of N-confident species to predict their invasiveness (Appendix S1: Table S3). Slightly more species grouped into very low and moderate invasiveness categories, and fewer in the low category, than with the Y-confident species, but an identical 38% of all N-confident species grouped into high or very high invasiveness states. Most of the model predictions for the N-confident species set were unambiguous, with strong probability outcomes of specific states, such as with *P. attenuata* predicted with 99.9% probability of high invasiveness. However, the model predictions of invasiveness of a few of the N-confident species straddled two or more invasiveness outcome states, although still with a dominant probability outcome. Examples are *P. chrysanthemi*, with the model predicting 52.7% probability of moderate invasiveness (the dominant outcome) and 47.3% probability of very low invasiveness, and *P. citricola*, with 48.2% moderate (dominant) and 47.7% very high invasiveness. Such bimodal outcome predictions, suggesting greater uncertainty in invasiveness response, occurred, however, with only 7 of the 101 N-confident species and resulted from their having uncommon combinations of covariate values such as for variables pertaining to measures of morphology or physiology as compared with the Y-confident species set on which the prediction model and its set of prior and conditional probabilities were constructed.

The invasiveness index was also calculated separately from the model for each N-confident species, and we found that the model-predicted values did not align well with the calculated values for this group (Figure 3). This is not unexpected. The invasiveness index is generated based on the number of countries and hosts recorded for a particular *Phytophthora* species, and for newly described species, very little information is available. Thus, it would be expected that the index would be underestimated as shown by the number of species with low invasiveness indices predicted by the Bayesian model to be in the high invasive categories (Appendix S1: Tables S2 and S3).

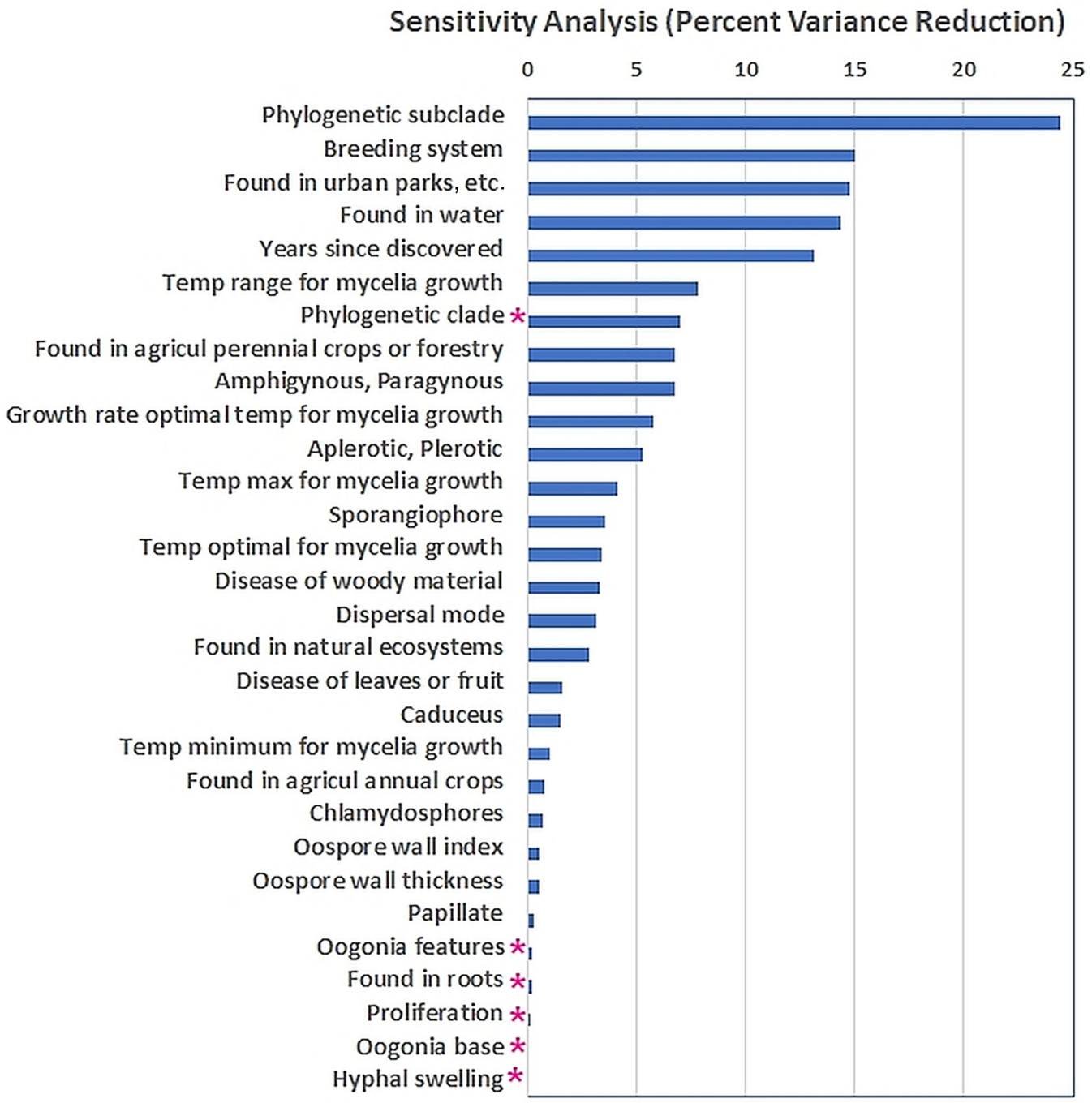


**FIGURE 1** Results of sensitivity analysis of the full model of predicting invasiveness of *Phytophthora* species with 30 covariates. An asterisk indicates a covariate identified as having the least sensitivity effect (lowest sensitivity values) or that is fully redundant with other variables (phylogenetic clade redundant with phylogenetic subclade) and subsequently removed in the final model. Agricul, agriculture; Temp, temperature.

Further, for long-described species now known to be part of species complexes, the country and host records may be overestimated and could result in a higher than expected invasiveness index value. However, in the cases examined here, the invasive categories of the long-described species were comparable with the invasiveness index except *P. pini*, which had a lower invasiveness index than was predicted; although described in 1925, this species was until recently considered a synonym for *P. citricola* and was only reinstated in 2010.

## DISCUSSION

Our modeling approach predicts the potential invasiveness of *Phytophthora* species by use of a broad suite of

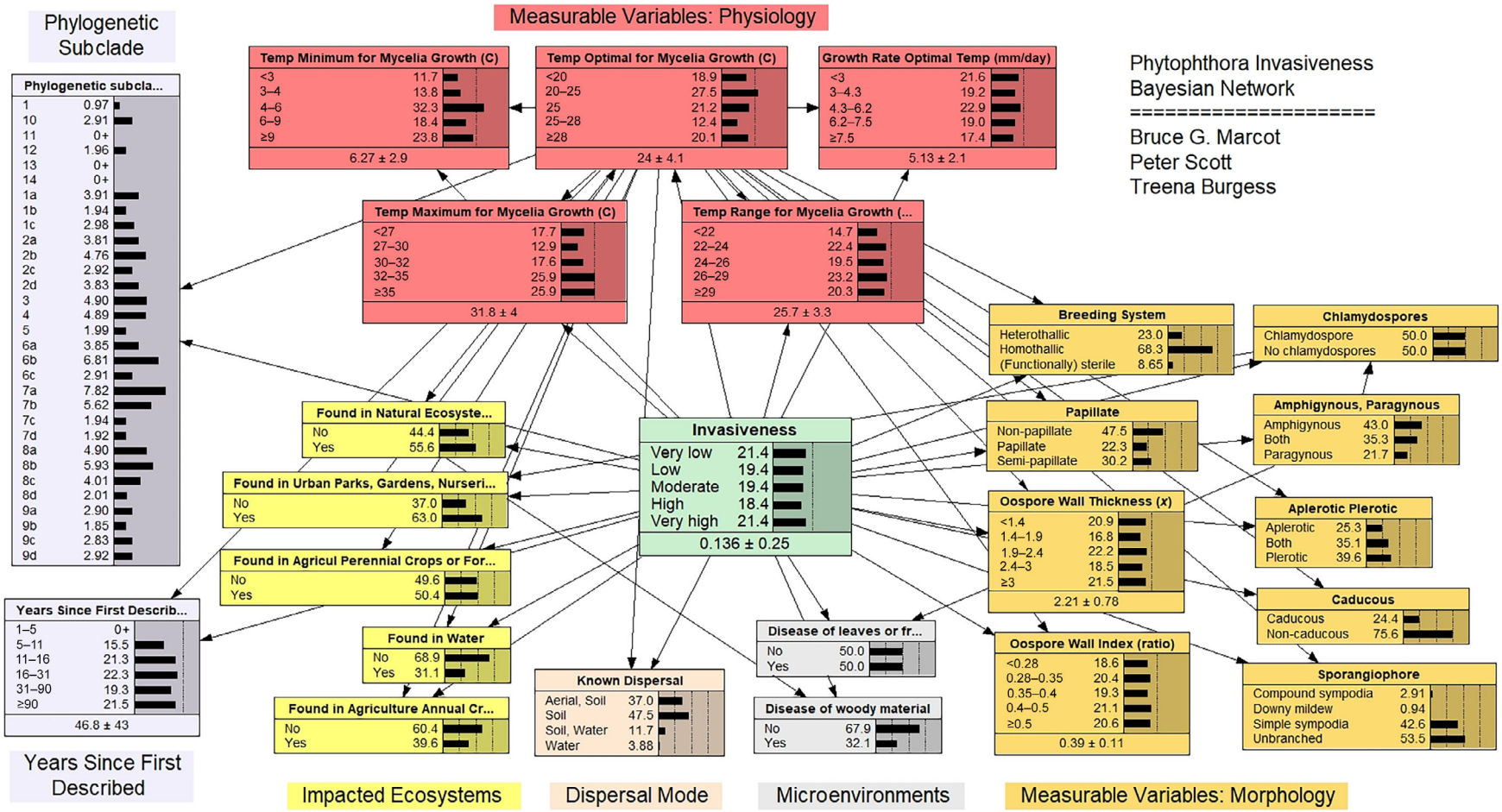


**FIGURE 2** Structure of the final Bayesian network model predicting invasiveness states of *Phytophthora* species. See Table 1 for full variable names. Agricul, agriculture; Temp, temperature.

**TABLE 4** Results of classification error rates and scoring rules comparing the initial full Bayesian network model and the final Bayesian network model.

| Metric | Initial full model | Final model |
|---|---|---|
| Classification error rate | 10.68% | 0.97% |
| Logarithmic loss | 0.235 | 0.052 |
| Quadratic loss | 0.153 | 0.019 |
| Spherical payoff | 0.915 | 0.990 |

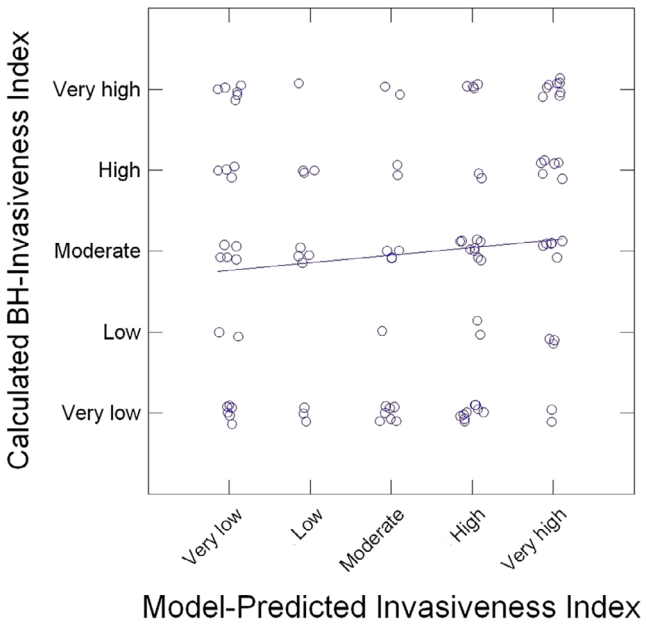


**FIGURE 3** Comparison of calculated branch height invasiveness index values against the model-predicted values for the N-confident *Phytophthora* species (see text for definition and Appendix S1: Tables S2 and S3 for species-specific values).

covariate descriptors organized into seven categories of taxonomy (clade and subclade), historical knowledge (years since first described), impacted ecosystems, microenvironments inhabited, dispersal mode, physiology, and morphology (Figure 2). The model was parameterized using the case file of species for which we have high confidence in their attributes and in their known invasiveness status. Invasiveness was calculated by measuring the branch height of a cladogram produced using measures of the number of host and country, and host and country identification rate for the different species.

Our novel approach using BN modeling thus is more comprehensive than other studies more focused on determining the establishment of selected *Phytophthora* species at specific locations or regions (e.g., Hernández-Lambraño et al., 2018; Redondo et al., 2018; Shearer & Crane, 2014) or studies determining specific determinants or modes of *Phytophthora* species dispersal (e.g., Krull et al., 2012; Martín-García et al., 2015). Also, the BN model can still provide results in the face of missing or incomplete data by defaulting to the prior probability distributions for input variables (as shown in Figure 2) that are based on the overall suite of conditions of the Y-confidence species set used to parameterize the model's probability structure.

## Key traits

The invasiveness index, which was first used by Scott et al. (2019) and utilized here to train the BN model parameters for the Y-confident group of species, uses the rate of spread of a species in countries and on hosts; however, these records are not necessarily definitive evidence of disease. Some of the most invasive species are well-known pathogens, but some water specialists are also invasive. It is important to know the invasiveness of these species because they may encounter a new susceptible host in a new environment, but the data generated for the N-confident group of species need to be interpreted with knowledge of the measured traits.

The model was most sensitive to the traits of phylogenetic subclade, breeding system, the microenvironment urban parks, found in water, years since discovered, and temperature range for mycelial growth. Phylogenetic subclade was the trait that best predicted invasiveness. This is not unexpected, as closely related species have a single common ancestor and share many traits; therefore, their phylogenetic placement will encompass much that is shared. Breeding system denotes whether a species is sterile (does not produce oospores), self-fertile (homothallic), or out-crossing (heterothallic). The BN model predicts that invasiveness of sterile species is more determined by other traits; that homothallic species have a 10% probability of very high and 30% of high or very high invasiveness; and that heterothallic species have a 55% probability of very high and 68% of high or very high invasiveness. Species in the same phylogenetic subclade usually have the same breeding system (Abad et al., 2023).

The number of years since a species has been described is also a strong predictor. This is expected, as long-described species tend to cause disease in agricultural commodities and have been reported from field and tree crops around the globe (Abad et al., 2023). Conversely, many newly described species have been found discovered during surveys of natural forests and may not have had the same opportunities for dispersal (Burgess et al., 2021). This does not mean, however, that they could not become significant pathogens if they were spread to a new environment and encounter naïve hosts. For example, our Bayesian analysis predicts that the undescribed species *Phytophthora* sp. *awatangi*, known

only from the highlands of Papua New Guinea, has the potential to be highly invasive. Another example is *P. kelmanii*, a species related to the widely distributed and well-known pathogen, *P. cryptogea*. *P. kelmanii* is also predicted to be highly invasive, and this is supported by the re-evaluation of GenBank data, which shows this species can already be found in 21 countries. On this basis, we retained time since discovery as a key input variable in the model, as it contributes significantly to explanatory power to predict invasiveness. However, if the value of this, or any, input variable for a given species is uncertain or unknown, its value does not have to be explicitly specified; instead, the generalized, prior probability distribution of that variable, based on the 103 known species used to parameterize the model, would be used. This would be the best use of the variable time since discovery for newly discovered species to retain the overall explanatory power of that information.

The environment in which a species is found also influenced its invasiveness with detection in urban environments, nurseries, perennial crops, and water being associated with a higher invasiveness. It has previously been demonstrated that urban environments and botanical gardens act as a bridgehead for *Phytophthora* invasions, providing a suitable environment for species to become established and then spread to newly encountered hosts (Hulbert et al., 2019; Paap et al., 2017). The number of species established in natural systems is more limited due to environmental filtering perhaps linked with healthier soils and more resilient environment (Dale et al., 2022; Redondo et al., 2018). The temperature range for mycelial growth and the growth rate at optimal temperature were also important traits influencing tolerance to temperature extremes and ability to rapidly colonize suitable substrate. Faster growing species would be able to colonize a host rapidly and competitively exclude other species (Sarker et al., 2021). Species with a narrow range will have a reduced climatic envelope and a restricted distribution such as species in clade 8b that cause disease in temperate crops (Barwell et al., 2021).

Some traits were less influential than may have been predicted. For example, oospores are dormant sexual propagules that persist in soil, in asymptomatic plants, and in plant debris (for reproduction and survival) (Dorrance, 2018; Gyeltshen et al., 2021). A high oospore wall index (the thickness of oospore walls relative to oospore volume) would be expected to improve survival for extreme temperatures and desiccation (survival) and hence survival during transport, and yet this was not a major predictor of invasiveness. The ability to infect the roots of plants was also a poor predictor of invasiveness; in this case, it may be because it is such a common trait for most species and as such does not discriminate among species.

## Model predictions as decision-aiding tools

Our ultimate aim in developing this invasiveness prediction model is to serve as a decision-aiding tool for local or regional managers engaged in trying to anticipate if any, and which, *Phytophthora* species might prove to be a critical pathogen. Our species database and model projections can help identify potentially impacted ecosystems and environments, dispersal media, and other environmental and organismic characteristics that may help focus preventative management activities in advance of pathogen spread and impact. Presenting outcomes as probabilities of invasiveness degrees is a way to express uncertainty of knowledge, which managers can use in their expression of risk attitude for their risk management decision-making.

Further, the database and model can be used to identify key species or habitat attributes potentially influencing invasiveness predictions but which are not well known, to help prioritize activities such as inventory, monitoring, and research to reduce those key uncertainties. Should establishment and spread already be suspected, some efficient survey techniques such as environmental DNA (eDNA) detection can be used (e.g., Hauck et al., 2021; Khaliq et al., 2018).

Some examples of local control of *Phytophthora* species include management of sudden oak death cited further above (also see Cunniffe et al., 2016; Filipe et al., 2012) and use of microbiota biocontrol agents to control the impact of *P. capsici* on agricultural products (Cucu et al., 2020). Other examples include reducing the spread of the soil pathogen *P. cinnamomi* during mining operations (Colquhoun & Kerp, 2007), its eradication from mine sites and sensitive ecosystems (Dunstan et al., 2010, 2020), and the use of phosphite to control its impact on trees (Crane & Shearer, 2014). Further examples can be found in the literature. Key to preventing infestations in the first place is knowing the potential for and mechanisms of, and designing the means to thwart, establishment and spread.

The model is intended to be refined with further knowledge and findings. To this end, we are providing access to the database and the model, especially, for management use, and also to encourage refinement of the data and updating of the model network and probability structure as provided by results of new studies.

## CONCLUSIONS

We have developed and provided a BN model predicting the potential degree of invasiveness of *Phytophthora* species. The model is intended to serve as a decision-aiding

tool to help identify conditions leading to potential establishment and spread of *Phytophthora* pathogens, and to be updated over time with new data and study findings.

## AUTHOR CONTRIBUTIONS

Bruce G. Marcot, Peter Scott, and Treena I. Burgess conceived the research, methods, and formal analysis, and performed data validation. Bruce G. Marcot developed and executed the Bayesian network model. Bruce G. Marcot and Peter Scott conducted data analysis. Peter Scott and Treena I. Burgess performed data curation. Bruce G. Marcot prepared the original draft manuscript. Peter Scott, Bruce G. Marcot, and Treena I. Burgess contributed to writing and editing. All authors read and approved the manuscript.

## ACKNOWLEDGMENTS

Bruce G. Marcot acknowledges support from the U.S. Forest Service, Pacific Northwest Research Station. Mention of commercial products does not necessarily imply endorsement by the U.S. Government. Peter Scott acknowledges support from the Department of Primary Industries and Regional Development, Perth, Western Australia. The Chief Executive Office of the Department of Primary Industries and Regions Development and the State of Western Australia accept no liability whatsoever by reason of negligence or otherwise arising from the use, or research of this information or any part of it. Peter Scott also acknowledges support from Nari Williams, Plant and Food Research, Havelock North New Zealand, and Martin Bader, Linnaeus University Sweden.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

## DATA AVAILABILITY STATEMENT

Data (Burgess & Scott, 2023) are available from Mendeley Data: https://doi.org/10.17632/6rshgjfr8w.1.

## ORCID

*Bruce G. Marcot* https://orcid.org/0000-0002-3667-7481

## REFERENCES

Abad, Z. G., T. Burgess, A. J. Redford, J. C. Bienapfl, R. Mathew, S. K. Srivastava, and K. C. Jennings. 2023. "Idphy: An International Online Resource for Molecular and Morphological Identification of *Phytophthora*." *Plant Disease* 107: 987–98.

Acebo-Guerrero, Y., A. Hernández-Rodríguez, M. Heydrich-Pérez, M. El Jaziri, and A. N. Hernández-Lauzardo. 2012. "Management of Black Pod Rot in Cacao (*Theobroma cacao* L.): A Review." *Fruits* 67: 41–8.

Barwell, L. J., A. Perez-Sierra, B. Henricot, A. Harris, T. I. Burgess, G. Hardy, P. Scott, et al. 2021. "Evolutionary Trait-Based Approaches for Predicting Future Global Impacts of Plant Pathogens in the Genus *Phytophthora*." *Journal of Applied Ecology* 58: 718–30.

Bourret, T. B., R. A. Choudhury, H. K. Mehl, C. L. Blomquist, N. McRoberts, and D. M. Rizzo. 2018. "Multiple Origins of Downy Mildews and Mito-Nuclear Discordance within the Paraphyletic Genus *Phytophthora*." *PLoS One* 13: e0192502.

Bradshaw, R. E., S. E. Bellgard, A. Black, B. R. Burns, M. L. Gerth, R. L. McDougal, P. M. Scott, et al. 2020. "*Phytophthora agathidicida*: Research Progress, Cultural Perspectives and Knowledge Gaps in the Control and Management of Kauri Dieback in New Zealand." *Plant Pathology* 69(1): 3–16.

Burgess, T., and P. Scott. 2023. "Multivariate Bayesian Analysis to Predict Invasiveness of *Phytophthora* Pathogens." Mendeley Data, V1. https://doi.org/10.17632/6rshgjfr8w.1.

Burgess, T. I., J. Edwards, A. Drenth, T. Massenbauer, J. Cunnington, R. Mostowfizadeh-Ghalamfarsa, Q. Dinh, et al. 2021. "Current Status of *Phytophthora* in Australia." *Persoonia* 47: 151–77.

Burgess, T. I., J. K. Scott, K. L. McDougall, M. J. C. Stukely, C. Crane, W. A. Dunstan, F. Brigg, et al. 2017. "Current and Projected Global Distribution of *Phytophthora cinnamomi*, One of the World's Worst Plant Pathogens." *Global Change Biology* 23(4): 1661–74.

Burgess, T. I., D. White, K. L. McDougall, J. Garnas, W. A. Dunstan, S. Català, A. J. Carnegie, et al. 2017. "Distribution and Diversity of *Phytophthora* across Australia." *Pacific Conservation Biology* 23: 150–62.

Cahill, D. M., J. E. Rookes, B. A. Wilson, L. Gibson, and K. L. McDougall. 2008. "*Phytophthora cinnamomi* and Australia's Biodiversity: Impacts, Predictions and Progress Towards Control." *Australian Journal of Botany* 56(4): 279–310.

Cardillo, E., A. Acedo, and E. Abad. 2018. "Topographic Effects on Dispersal Patterns of *Phytophthora cinnamomi* at a Stand Scale in a Spanish Heathland." *PLoS One* 13(3): e0195060.

Cayuela, L., A. Stein, and J. Oksane. 2021. "Taxonstand: Taxonomic Standardization of Plant Species Names." R Package Version 2.3. https://cran.r-project.org/package=Taxonstand.

Cobb, R. C., M. N. Chan, R. K. Meentemeyer, and D. M. Rizzo. 2012. "Common Factors Drive Disease and Coarse Woody Debris Dynamics in Forests Impacted by Sudden Oak Death." *Ecosystems* 15(2): 242–55.

Colquhoun, I. J., and N. L. Kerp. 2007. "Minimizing the Spread of a Soil-Borne Plant Pathogen during a Large-Scale Mining Operation." *Restoration Ecology* 15(4): S85–93.

Cooke, L. R., H. T. A. M. Schepers, A. Hermansen, R. A. Bain, N. J. Bradshaw, F. Ritchie, D. S. Shaw, et al. 2011. "Epidemiology and Integrated Control of Potato Late Blight in Europe." *Potato Research* 54(2): 183–222.

Crane, C. E., and B. L. Shearer. 2014. "Comparison of Phosphite Application Methods for Control of *Phytophthora cinnamomi* in Threatened Communities." *Australasian Plant Pathology* 43: 143–9.

Cucu, M. A., G. Gilardi, M. Pugliese, I. Ferrocino, and M. L. Gullino. 2020. "Effects of Biocontrol Agents and Compost against the *Phytophthora capsici* of Zucchini and Their Impact

on the Rhizosphere Microbiota." *Applied Soil Ecology* 154: 103659.

Cunniffe, N. J., R. C. Cobb, R. K. Meentemeyer, D. M. Rizzo, and C. A. Gilligan. 2016. "Modeling When, Where, and How to Manage a Forest Epidemic, Motivated by Sudden Oak Death in California." *PNAS* 113(20): 5640–5.

Dale, A. L., N. Feau, J. A. Berube, J. Ponchart, G. J. Bilodeau, and R. C. Hamelin. 2022. "Urban Environments Harbor Greater Oomycete and *Phytophthora* Diversity, Creating a Bridgehead for Potential New Pathogens to Natural Ecosystems." *Environmental DNA* 4: 1039–51.

Davidson, J. M., A. C. Wickland, H. A. Patterson, K. R. Falk, and D. M. Rizzo. 2005. "Transmission of *Phytophthora ramorum* in Mixed-Evergreen Forest in California." *Ecology and Epidemiology* 95: 587–96.

Dillon, W. W., and R. K. Meentemeyer. 2019. "Direct and Indirect Effects of Forest Microclimate on Pathogen Spillover." *Ecology* 100(5): e02686.

Dlamini, W. M. 2011. "A Data Mining Approach to Predictive Vegetation Mapping Using Probabilistic Graphical Models." *Ecological Informatics* 6: 111–24.

Dorrance, A. E. 2018. "Management of *Phytophthora sojae* of Soybean: A Review and Future Perspectives." *Canadian Journal of Plant Pathology* 40(2): 210–9.

Dunstan, W. A., K. Howard, A. Grigg, C. Shaw, T. I. Burgess, and G. E. S. Hardy. 2020. "Towards Eradication of *Phytophthora cinnamomi* Using a Fallow Approach in a Mediterranean Climate." *Forests* 11(10): 1101.

Dunstan, W. A., T. Rudman, B. L. Shearer, N. A. Moore, T. Paap, M. C. Calver, B. Dell, and G. E. S. J. Hardy. 2010. "Containment and Spot Eradication of a Highly Destructive, Invasive Plant Pathogen (*Phytophthora cinnamomi*) in Natural Ecosystems." *Biological Invasions* 12(4): 913–25.

Farr, D. F., and A. Y. Rossman. 2022. "Fungal Databases, U.S. National Fungus Collections, ARS, USDA." https://nt.ars-grin.gov/fungaldatabases/.

Fernández-Habas, J., P. Fernández-Rebollo, M. R. Casado, A. M. G. Moreno, and B. Abellanas. 2019. "Spatio-Temporal Analysis of Oak Decline Process in Open Woodlands: A Case Study in SW Spain." *Journal of Environmental Management* 248: 109308.

Filipe, J. A. N., R. C. Cobb, R. K. Meentemeyer, C. A. Lee, Y. S. Valachovic, A. R. Cook, D. M. Rizzo, and C. A. Gilligan. 2012. "Landscape Epidemiology and Control of Pathogens with Cryptic and Long-Distance Dispersal: Sudden Oak Death in Northern Californian Forests." *PLoS Computational Biology* 8(1): e1002328.

Galili, T. 2015. "dendextend: An R Package for Visualizing, Adjusting and Comparing Trees of Hierarchical Clustering." *Bioinformatics* 31(22): 3718–20.

Garrett, K. A., L. N. Zúñiga, G. A. Forbes, C. C. Mundt, Z. Su, and R. J. Nelson. 2009. "Intraspecific Functional Diversity in Hosts and Its Effect on Disease Risk across a Climatic Gradient." *Ecological Applications* 19(7): 1868–83.

Gaydos, D. A., A. Petrasolva, R. C. Cobb, and R. K. Meentemeyer. 2019. "Forecasting and Control of Emerging Infectious Forest Disease through Participatory Modelling." *Philosophical Transactions of the Royal Society B* 374: 20180283.

Graham, J., and E. Feichtenberger. 2015. "Citrus *Phytophthora* Diseases: Management Challenges and Successes." *Journal of Citrus Pathology* 2(1): iocv_journalcitruspathology_27203.

Granke, L. L., L. Quesada-Ocampo, K. Lamour, and M. K. Hausbeck. 2012. "Advances in Research on *Phytophthora capsici* on Vegetable Crops in the United States." *Plant Disease* 96(11): 1588–600.

Gyeltshen, J., W. A. Dunstan, A. H. Grigg, T. I. Burgess, and G. E. St. J. Hardy. 2021. "The Influence of Time, Soil Moisture and Exogenous Factors on the Survival Potential of Oospores and Chlamydospores of *Phytophthora cinnamomi*." *Forest Pathology* 51(1): e12637.

Hansen, E. M. 2015. "*Phytophthora* Species Emerging as Pathogens of Forest Trees." *Current Forestry Reports* 1: 16–24.

Hauck, L. L., K. A. Weitemier, B. E. Penaluna, T. S. Garcia, and R. Cronn. 2021. "Casting a Broader Net: Using Microfluidic Metagenomics to Capture Aquatic Biodiversity Data from Diverse Taxonomic Targets." *Environmental DNA* 1: 251–67.

Hernández-Lambraño, R. E., P. González-Moreno, and J. Á. Sánchez-Agudo. 2018. "Environmental Factors Associated with the Spatial Distribution of Invasive Plant Pathogens in the Iberian Peninsula: The Case of *Phytophthora cinnamomi* Rands." *Forest Ecology and Management* 419–420: 101–9.

Hulbert, J. M., T. Paap, T. I. Burgess, R. Roets, and M. J. Wingfield. 2019. "Botanical Gardens Provide Baseline *Phytophthora* Diversity Data." *Urban Forestry and Urban Greening* 46: e126461.

Ireland, K. B., G. E. S. J. Hardy, and D. J. Kriticos. 2013. "Combining Inferential and Deductive Approaches to Estimate the Potential Geographical Range of the Invasive Plant Pathogen, *Phytophthora ramorum*." *PLoS One* 8(5): e63508.

Jiang, L., Z. Cai, D. Wang, and H. Zhang. 2012. "Improving Tree Augmented Naive Bayes for Class Probability Estimation." *Knowledge-Based Systems* 26: 239–45.

Khaliq, I., G. E. J. St, D. W. Hardy, and T. I. Burgess. 2018. "eDNA from Roots: A Robust Tool for Determining *Phytophthora* Communities in Natural Ecosystems." *FEMS Microbiology Ecology* 94(5): fiy048.

Kindt, R. 2021. "WorldFlora: An R Package for Exact and Fuzzy Matching of Plant Names against the World Flora Online Taxonomic Backbone Data." *Applications in Plant Sciences* 8(9): e11388.

Krah, F.-S. 2016. "rUSDA: Interface to USDA Databases." R Package Version 1.0.7. https://CRAN.R-project.org/package=rUSDA.

Krull, C. R., N. W. Waipara, D. Choquenot, B. R. Burns, A. M. Gormley, and M. C. Stanley. 2012. "Absence of Evidence Is Not Evidence of Absence: Feral Pigs as Vectors of Soil-Borne Pathogens." *Austral Ecology* 38(5): 534–42.

Laliberté, E., H. Lambers, T. I. Burgess, and S. J. Wright. 2015. "Phosphorus Limitation, Soil-Borne Pathogens and the Coexistence of Plant Species in Hyperdiverse Forests and Shrublands." *New Phytologist* 206(2): 507–21.

Marano, A. V., A. L. Jesus, J. I. de Souza, G. H. Jerônimo, D. R. Gonçalves, M. C. Boro, S. C. O. Rocha, and C. L. A. Pires-Zottarelli. 2015. "Ecological Roles of Saprotrophic *Peronosporales* (Oomycetes, Straminipila) in Natural Environments." *Fungal Ecology* 19: 77–88.

Marcot, B. G. 2006. "Characterizing Species at Risk I: Modeling Rare Species under the Northwest Forest Plan." *Ecology and Society* 11(2): 10.

Marcot, B. G. 2012. "Metrics for Evaluating Performance and Uncertainty of Bayesian Network Models." *Ecological Modelling* 230: 50–62.

Marcot, B. G., and A. Hanea. 2021. "What Is an Optimal Value of K in K-Fold Cross-Validation in Discrete Bayesian Network Analysis?" *Computational Statistics* 36(3): 2009–31.

Martín-García, J., A. Solla, T. Corcobado, E. Siasou, and S. Woodward. 2015. "Influence of Temperature on Germination of *Quercus ilex* in *Phytophthora cinnamomi*, *P. gonapodyides*, *P. quercina* and *P. psychrophila* Infested Soils." *Forest Pathology* 45(3): 215–23.

McCann, R., B. G. Marcot, and R. Ellis. 2006. "Bayesian Belief Networks: Applications in Natural Resource Management." *Canadian Journal of Forest Research* 36(12): 3053–62.

Meentemeyer, R. K., N. J. Cunniffe, A. R. Cook, J. A. N. Filipe, R. D. Hunter, D. M. Rizzo, and C. A. Gilligan. 2011. "Epidemiological Modeling of Invasion in Heterogeneous Landscapes: Spread of Sudden Oak Death in California (1990–2030)." *Ecosphere* 2(2): 17.

Meurisse, N., B. G. Marcot, O. Woodberry, B. Barratt, and J. Todd. 2021. "Risk Analysis Frameworks Used in Biological Control and Introduction of a Novel Bayesian Network Tool." *Risk Analysis* 42(6): 1255–76.

Nagel, J. H., M. Gryzenhout, B. Slippers, and M. J. Wingfield. 2013. "The Occurrence and Impact of *Phytophthora* on the African Continent." In *Phytophthora, a Global Perspective*, edited by K. Lamour. Boston, MA: CAB International.

NRC. 2012. *Assessing the Reliability of Complex Models Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification*. Washington, DC: National Research Council, The National Academies Press. 144 pp.

Paap, T., T. I. Burgess, and M. J. Wingfield. 2017. "Urban Trees: Bridge-Heads for Forest Pest Invasions and Sentinels for Early Detection." *Biological Invasions* 19: 3515–26.

Panth, M., S. C. Hassler, and F. Baysal-Gurel. 2020. "Methods for Management of Soilborne Diseases in Crop Production." *Agriculture* 10(1): 16. https://doi.org/10.3390/agriculture10010016.

Parke, J., E. Mittermaier, N. Redekar, J. Eberhart, and T. Matteson. 2020. "A Survey of *Phytophthora*, *Pythium*, and *Phytopythium* Species in Soil from Upland Prairie Restoration Sites in Western Oregon." In *Proceedings of the Seventh Sudden Oak Death Science and Management Symposium: Healthy Plants in a World with Phytophthora*. USDA Forest Service, General Technical Report PSW-GTR-268, edited by S. J. Frankel and J. M. Alexander, 91–3. Albany, CA: Pacific Southwest Research Station.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.

Ramírez-Gil, J. G., D. A. Castañeda-Sánchez, and J. G. Morales-Osorio. 2017. "Production of Avocado Trees Infected with *Phytophthora cinnamomi* under Different Management Regimes." *Plant Pathology* 66(4): 623–32.

Redondo, M. A., J. Boberg, J. Stenlid, and J. Oliva. 2018. "Functional Traits Associated with the Establishment of Introduced *Phytophthora* spp. in Swedish Forests." *Journal of Applied Ecology* 55(3): 1538–52.

Rizzo, D. M., and M. Garbelotto. 2003. "Sudden Oak Death: Endangering California and Oregon Forest Ecosystems." *Frontiers in Ecology and the Environment* 1(5): 197–204.

Ruffner, B., D. Rigling, and C. A. Schoebel. 2019. "Multispecies *Phytophthora* Disease Patterns in Declining Beech Stands." *Forest Pathology* 49: e12514.

Sarker, S. R., J. McComb, T. I. Burgess, and G. E. S. J. Hardy. 2021. "Timing and Abundance of Sporangia Production and Zoospore Release Influences the Recovery of Different *Phytophthora* Species by Baiting." *Fungal Biology* 125(6): 477–84.

Schoebel, C. N., J. Stewart, N. J. Gruenwald, D. Rigling, and S. Prospero. 2014. "Population History and Pathways of Spread of the Plant Pathogen *Phytophthora plurivora*." *PLoS One* 9(1): e85368.

Scott, P., M. K.-F. Bader, T. Burgess, G. Hardy, and N. Williams. 2019. "Global Biogeography and Invasion Risk of the Plant Pathogen Genus *Phytophthora*." *Environmental Science & Policy* 101: 175–82.

Shearer, B. L., and C. E. Crane. 2014. "*Phytophthora cinnamomi* Disease Expression and Habitat Suitability of Soils on a Topographic Gradient across a Coastal Plain from Dunes to Forested Peneplain." *Australasian Plant Pathology* 43: 131–42.

Tsykun, T., S. Prospero, C. N. Schoebel, A. Rea, and T. I. Burgess. 2022. "Global Invasion History of the Emerging Plant Pathogen *Phytophthora multivora*." *BMC Genomics* 23: 153.

Václavík, T., and R. K. Meentemeyer. 2009. "Invasive Species Distribution Modeling (iSDM): Are Absence Data and Dispersal Constraints Needed to Predict Actual Distributions?" *Ecological Modelling* 220(23): 3248–58.

Warnes, M. G. R., B. Bolker, L. Bonebakker, R. Gentleman, and W. Huber. 2022. "Package 'gplots'." Various R Programming Tools for Plotting Data. https://cran.r-project.org/web/packages/gplots/gplots.pdf.

Wesner, J. S., and J. P. F. Pomeranz. 2021. "Choosing Priors in Bayesian Ecological Models by Simulating from the Prior Predictive Distribution." *Ecosphere* 12(9): e03739.

Yang, X., B. M. Tyler, and C. Hong. 2017. "An Expanded Phylogeny for the Genus *Phytophthora*." *IMA Fungus* 8: 355–84.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Marcot, Bruce G., Peter Scott, and Treena I. Burgess. 2023. "Multivariate Bayesian Analysis to Predict Invasiveness of *Phytophthora* Pathogens." *Ecosphere* 14(6): e4573. https://doi.org/10.1002/ecs2.4573