Modeling when, where, and how to manage a forest epidemic, motivated by sudden oak death in California

Nik J. Cunniffea, Richard C. Cobbb, Ross K. Meentemeyer, David M. Rizzo, and Christopher A. Gilligan

Introduction

Sudden oak death, caused by Phytophthora ramorum, has killed millions of oak and tanoak in California since its first detection in 1995. Despite some localized small-scale management, there has been no large-scale attempt to slow the spread of the pathogen in California. Here we use a stochastic spatially explicit model parameterized using data on the spread of P. ramorum to investigate whether and how the epidemic can be controlled. We find that slowing the spread of P. ramorum is now not possible, and has been impossible for a number of years. However, despite extensive cryptic (i.e., presymptomatic) infection and frequent long-range transmission, effective exclusion of the pathogen from large parts of the state could, in principle, have been possible were it to have been started by 2002. This is the approximate date by which sufficient knowledge of P. ramorum epidemiology had accumulated for large-scale management to be realistic. The necessary expenditure would have been very large, but could have been greatly reduced by optimizing the radius within which infected sites are treated and careful selection of sites to treat. In particular, we find that a dynamic strategy treating sites on the epidemic wave front leads to optimal performance. We also find that “front loading” the budget, that is, treating very heavily at the start of the management program, would greatly improve control. Our work introduces a framework for quantifying the likelihood of success and risks of failure of management that can be applied to invading pests and pathogens threatening forests worldwide.

Phytophthora ramorum | constrained budget | landscape-scale stochastic epidemiological model | optimizing disease control | risk aversion

Introductions of new pathogens into previously uncolonized areas pose threats to trees in natural ecosystems, commercial woodlands, and urban environments. Rates of introduction are increasing (1), driven by changing climate (2) and altered patterns of travel and trade (3). Emerging epidemics cause direct economic loss from death and restricted growth of trees grown for timber and horticultural use (4). Other major impacts occur when susceptible trees play critical roles in ecosystem services (5).

Successful control of epidemics involves matching the scale of management with the inherent scale of spread (6). Early detection and timely removal of affected trees from a small number of newly infected sites soon after initial introduction(s) can prevent epidemics (7). Routine detection and control of emerging pathogens of woodland trees, however, are prone to significant logistical and epidemiological constraints. Detection and reporting can be delayed by incomplete and infrequent sampling of large areas of susceptible hosts (8), with broad pathogen host ranges often increasing the area that must be surveyed. Infected sites may be inaccessible or under multiple ownership (9). Long incubation periods for some pathogens mean that disease remains cryptic while infection continues to spread (10). Spread can also be over long distances (11), with extensive creation of new foci.

Difficulties in detection and management have undoubtedly contributed to high-profile failures of large-scale control programs for a number of tree diseases, including chestnut blight (12), white pine blister rust (13), Dutch elm disease (10), and citrus canker (14). Identifying when, where, why, and how to manage emerging epidemics at regional, state, or countrywide scales, and even whether or not it is feasible to do so, remains a major challenge (15). However, understanding whether management can eradicate a pathogen or restrict its spread to uninvaded locations is critical in identifying cost-effective control strategies. We show here how mathematical modeling can be used to do this, using sudden oak death (SOD) in California (CA) as an example.

SOD, caused by the oomycete Phytophthora ramorum (PR), has killed millions of oak (Quercus spp.) and tanoak (Notholithocarpus densiflorus) in CA since first detection in 1995 (16). The epidemic has been intensively monitored (17), and much is now known about PR epidemiology. However, questions remain about the feasibility of statewide control, introducing uncertainty and confusion into identifying regional management objectives. The epidemic also provides an opportunity for retrospective analyses of how effective control scenarios could have been, had they been introduced at different stages in the epidemic.

Here we extend a previously tested, spatially explicit, stochastic statewide model, resolved to 250 × 250-m resolution (18), to compare the range of outcomes for different management scenarios, addressing the following questions:

Could statewide prevention of continued pathogen spread be successful were it to start now, given the current size of the epidemic and the budget potentially available for treatment?

Could prevention of pathogen spread ever have been successful by management starting after the pathogen was sufficiently well characterized for control to have been realistic?

How could local and statewide deployment of management have been optimized?

The analyses also address the following generic questions about epidemic control under uncertainty:

Significance

We use sudden oak death in California to illustrate how mathematical modeling can be used to optimize control of established epidemics of invading pathogens in complex heterogeneous landscapes. We use our statewide model—which has been parameterized to pathogen spread data—to address a number of broadly applicable questions. How quickly must management start? When is an epidemic too large to prevent further spread effectively? How should local treatment be deployed? How does this depend on the budget and level of risk aversion? Where should treatment be targeted? How should expenditure be balanced on detection and treatment? What if the budget changes over time? The underlying principles are important for management of all plant disease epidemics in natural ecosystems.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

1To whom correspondence should be addressed. Email: njc1001@cam.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1602153113/-/DCSupplemental.
How soon must control start for it to be effective?
When is an epidemic so large that control is impossible?
Which sites should be targeted when there are insufficient resources to treat all infected sites?
How extensively should sites be treated?
How does this depend on the budget and risk aversion?
How can costs of detection and treatment be balanced?
What is the effect of a budget that changes over time?

Over 100 tree and forest shrub species are known to be susceptible to PR infection (16). First reported in 1995 in coastal regions near San Francisco (19), the pathogen is transmitted locally via rain splash and over long distances via storms and human-mediated transport. Subsequent estimates of the initial introduction to natural ecosystems suggest that the first invasion occurred around 1990 (16). PR has spread widely in coastal CA since then: Coastal forests from southern Monterey County up to northern Humboldt County are affected. Billions of tanoak and oak trees, covering over 20 million ha of land, are potentially threatened (20). There are also large and growing epidemics in several countries in Europe (21).

The only treatment shown to be effective in reducing pathogen prevalence at the landscape scale is removal of host species (22), as has been practiced in the United Kingdom for a number of years (21). In North America, however, large-scale management has not been attempted. Nevertheless, an outbreak of SOD in Curry County, Oregon, has remained relatively small in comparison with CA due to active management by host removal since first detection in 2001 (23). The attempted eradication of PR in Oregon in the very earliest stage of that outbreak was only possible because the potential impacts were already clear from CA, and because its epidemiology was beginning to be characterized (22). We consider 2002 to be the earliest that statewide attempts to prevent further spread of PR in CA could, in principle, have been attempted. This was 1 y after the pathogen was first named (24), and the year in which the pathogen first came under European Community emergency control measures (21).

Results
Pathogen Spread Without Management. Our analyses predict that spread of PR within CA will accelerate after 2020 if measures to slow the spread are not attempted. This will be driven by the pathogen reaching the northwestern coast, where large regions of continuous host and suitable weather conditions facilitate spread (18). The estimated area infected by 2030 in CA is around 14,000 km², a nearly 10-fold increase from the predicted ~1,550 km² infected in 2014 (Fig. 1B and Movie S1). However, all estimates of future epidemic size reflect the inherent variability of pathogen spread (Fig. 1C), largely driven by the timing of long-distance dispersal to the northwest but also reflecting stochastic variability in pathogen bulking up and dispersal driven by fluctuating environmental drivers in a heterogeneous host landscape. A 95% prediction interval for the area infected by 2030 ranges from ~7,600 to ~19,600 km². Acceleration in spread means that an ever-increasing area is expected to be infected but in which the pathogen remains undetected. By 2030, we predict on average ~4,000 km² would have been colonized too recently to show crown mortality symptoms detectable via an aerial survey.

Slow the Spread Management Starting 2014. Initially, we consider performance of a baseline management strategy that removes PR-susceptible and -infected hosts within a radius of 375 m of detected loci of infection. Treated sites are randomly selected from the set of sites known to contain infection until a fixed budget per y is exhausted. Control starting in 2014 under this scenario has very little effect on the area lost by 2030, that is, the area ever affected by either disease or host removal (Fig. 2 and Movie S2). Even a very large budget allowing up to 200 km² to be treated annually—which we estimate would cost at least 100 million US dollars (USD) per y (SI Appendix)—has almost no effect on spread. The California Department of Forestry and Fire Protection (Cal Fire) allocated just over 90 million USD in fiscal year 2015–2016 to implement drought mitigation, suggesting this level of expense is at least within the bounds of possibility. However, multiple and sometimes competing objectives must be addressed by Cal Fire, meaning that allocation of the full budget to slowing the spread of a single pathogen is very unlikely. Given the almost imperceptible effect on the epidemic of even such a large amount of management, we consider prevention of spread at the statewide scale to no longer be possible.

Slow the Spread Management Starting Before 2014. Were an attempt to restrict the natural spread of PR to have started earlier than 2014, it could have been more successful, although the requisite budget increases rapidly as the start date becomes later (Fig. 3). Successful management can make it impossible to spend the full budget as the epidemic comes under control, and so the total area managed by 2030, which corresponds to the amount spent on treatment, is nonmonotone in budget and starting year (Fig. 3B). The distribution of area lost (Fig. 3C) again reveals the wide variability in outcome, a pattern that is replicated for the full range of budgets and starting years we considered.

Optimizing the Radius of Treatment. We consider how performance of management is affected by the treatment radius within which susceptible and infected hosts are removed around each detected site, focusing initially on a default budget that allows removal of up to 50 km²/y. Unless otherwise stated, henceforth all results relate to management starting in 2002, the first date we consider control to have been realistic. Too small a treatment radius does
not account sufficiently for cryptic infection around each detected focus, whereas at too large a treatment radius, too many healthy trees are unnecessarily removed and/or an increased proportion of other infection remains untreated because of the limited budget (Fig. 4A). There is therefore an optimum in the area lost as the treatment radius is altered, at least if the objective is phrased in terms of minimizing the average epidemic size. In particular, at the optimal radius of 187.5 m, the median area lost is ~6,800 km² (95% prediction interval ~2,600 to ~16,200 km²; compare median of ~7,900 km² with 95% interval ~1,100 to ~17,100 km² at the 375-m baseline).

**Effect of Budget and Risk Aversion.** The optimal treatment radius depends upon the budget. The more limited the resources, the more strongly treatment should focus on locations known to be infected by selecting a shorter treatment radius around each detected site, because this allows a larger number of distinct foci to be managed (Fig. 4B). Optimizing the treatment radius at different percentiles of the distribution of area lost reveals that shorter radii are also promoted when there is a high level of risk aversion (Fig. 4C). Again, a smaller treatment radius means that more known disease foci will be removed, which reduces the risk of failing to treat a site that goes on to cause many secondary infections.

**Optimizing the Subset of Locations to Treat.** We have assumed that foci to treat are selected randomly from the set of sites known to contain currently uncontrolled infection, analogous to independent action by stakeholders throughout infected areas. The selection of sites to treat can be improved (25–28). Logistical constraints surrounding movement of machinery and access to land make tackling highly infected regions an attractive default. However, this “infected” strategy (Movie S3) performs worse in terms of epidemic impact than the default “random” strategy (Fig. 4A and Movie S4). We considered a range of other possibilities, which in order of performance were focusing control on regions with large areas of host, irrespective of disease status (“host” strategy) (Movie S5), with high local rates of spread (“hazard” strategy (Movie S6), targeting areas with high basic reproductive number; R₀ (29)), or with large areas currently uninfected (“susceptible” strategy) (Movie S7).

However, the best-performing strategy we tested targets local control at and ahead of the northerly wave front of epidemic spread (“wave-front” strategy; Fig. 5A and Movie S8). This strategy reduces predicted area lost by 2030 to ~2,400 km² (95% interval 2,600–7,400 km²) at an optimal radius of 362.5 m (compare ~6,800 km² with interval 2,600–16,200 km² at the optimal radius 187.5 m under the random strategy starting 2002) (Fig. 5A and D). The increase in optimal treatment radius for the better-performing strategies was consistent across all priorizations we tested. Selecting sites for management more effectively allows more extensive treatment around each one, because sites that have a greater impact are then treated, reducing the future spread of the epidemic.

The relative performance at the optimal treatment radius of the six strategies was consistent across all control budgets (Fig. 5A and D).
5B) and starting dates (Fig. 5C) we considered, emphasizing that the ordering of these strategies is generic, at least for this epidemic. We note, however, that the performance of even the wave-front strategy degrades for control that starts later in the epidemic. Management starting in 2014 of up to 50 km²/y leads to an average area lost of greater than 7,000 km² by 2030. The epidemic remains out of control even when treating more effectively, and the ostensibly large reduction in area lost under the wave-front strategy starting in 2014 corresponds only to a 7-y delay relative to not treating at all (SI Appendix, Fig. S8). Such poor performance even with the optimal strategy and comparatively large budget (equivalent to at least 25 million USD per y) reiterates our earlier contention that the epidemic is uncontrollable at the statewide scale for management starting today.

**Budgets That Vary over Time.** Resources that can be devoted to control are set by policymakers in response to a number of complex and time-dependent drivers, including public opinion, other demands on a limited budget for plant health, and the perceived and anticipated success of control. We therefore assessed management strategies for which the budget varies over time but that ensure the total budget over the period 2002–2030 is fixed (SI Appendix, Fig. S4). It is better to devote larger resources to treatment early in the epidemic, particularly if any unused budget can be carried over to subsequent years, because the net rate of growth of the epidemic increases with its size, and so earlier treatment reduces the future growth rate (28).

**Accounting for the Cost of Detection.** Limited resources for management must be split between pathogen detection and treatment (30). Results thus far are contingent on extensive statewide surveys, repeated yearly. In practice, the cost of this would reduce the amount remaining to be spent on treatment. We therefore tested the effect of assuming the area that can be surveyed for disease symptoms is proportional to the fraction of the budget that is spent on detection, with the remainder devoted to management (SI Appendix, Fig. S1B). We find an optimum irrespective of the year treatment starts, and that a larger proportion of the budget should be devoted to detection for interventions that start earlier, because smaller amounts of treatment are then required and detection is more difficult when the epidemic is smaller.

**Discussion**
Deciding whether and how to control invasive pathogens are two of the principal challenges in epidemiology (15). Our results show how a stochastic, spatially explicit, epidemiological model can be used to integrate the current state of knowledge about the pathogen with detailed information on host topology and environmental drivers to predict the likely effects of management strategies. The model simulates epidemic spread through a large heterogeneous host landscape. It couples fine spatial resolution of 250 × 250 m for pathogen transmission with the facility to address local and statewide control of the advancing disease, allowing for spatial targeting of control within a limited budget. The model has been parameterized using data for PR spread in CA and successfully used to predict patterns of statewide spread (18). However, this is the first study, to our knowledge, to use a plant disease model calibrated to pathogen spread data and that predicts spread over such a large spatial scale to understand management. A version of the model has also been calibrated and validated for spread in the United Kingdom, in which context it is being used by plant health policymakers to inform management strategies, particularly the extent of host removal around infected larch stands (31).

We distinguish three phases of invasion and spread for an emerging epidemic.

Phase 1: initial invasion that may go undetected and undiagnosed for some time and in some places; frequently, little is known about the causal agent and the potential for damage.

Phase 2: the epidemic continues and is perceived to be a potentially serious threat; knowledge about the causal agent is sparse but there are sufficient data to construct models to begin to assess different strategies.

Phase 3: the pathogen has spread far enough that eradication is no longer possible; local containment may still be an option.

Our analyses show that statewide action to eradicate or even slow the spread of PR is no longer feasible, even with a substantial budget, indicating that the epidemic in CA is now firmly in phase 3 (Figs. 1 and 2). Shifting management resources to restoring degraded forests and protecting ecological function at smaller scales would be more beneficial than attempting statewide control. Prevention of spread could, however, have been feasible starting earlier, with substantial losses prevented by earlier treatment (Fig. 3 A and B), indicating that the epidemic was still in phase 2 in 2002. However, the cost in 2002 would have been very high, and practical implementation would have required unprecedented cooperation among agencies and landowners.

The analyses for SOD in CA illustrate how epidemiological principles could be translated into practical application. One important principle is matching the scale of treatment with the inherent spatial and temporal scales of pathogen spread to achieve effective management (6, 15, 25–27, 32). Characterization of epidemic scales is complicated, as the epidemic advances through heterogeneous host populations subject to variability in...
environmental drivers. Nevertheless, optimal scenarios can be derived for the median response, for example, the treatment radius that minimizes the impact of disease. Although optimal control radii have been derived using models at small spatial scales (7, 33), this is the first demonstration, to our knowledge, that the idea extends to landscape-scale control of plant disease within a fixed budget. Impact is assessed by totaling the area lost from disease and from removal of healthy trees around infected sites. However, the metric could readily be extended to a range of objective functions with different weightings for the individual components depending on perceived costs and benefits by different stakeholders (7, 34). Treatment radii can also be adjusted to allow for different degrees of risk aversion (35), for example, selecting the radius that corresponds to the 5th percentile (high risk aversion) through to the 95th percentile (low risk aversion) (compare Fig. 4).

Our results confirm that it would have been possible, in principle, to bring the epidemic under control by early removal of infected hosts. In practice, however, selection of sites for treatment within a limited budget is problematic. We considered a range of possibilities from random selection of sites for treatment, analogous to independent action by stakeholders in infected areas, through to highly structured centralized selection of sites according to model predictions (28, 36), the first assessment, to our knowledge, of such a prioritization for a plant disease. Model-informed selection proved superior to randomly choosing sites in terms of minimizing areas lost. The impacts of disease and treatment successively decreased (Fig. 5) by prioritizing areas where there are large host populations, where the potential for rapid local spread is high [i.e., selecting regions of high hazard (15, 18, 29)], and where the density of susceptible hosts available for infection is high. However, for the combination of host distribution, pathogen dispersal, epidemic progress, and environmental suitability for SOD in CA, local control at and ahead of the northerly wave front proved most effective (Fig. 5).

Finding and mapping newly infected sites are expensive for an epidemic spreading through a heterogeneous host landscape in which access is limited by terrain, private ownership, and resources (8). Previous work has shown how to optimize the balance between detection and treatment to maximize the cost-effectiveness when there is a shared budget for detection and treatment, but has focused exclusively on simple theoretical models (37, 38). Here we present the first test, to our knowledge, of the tradeoff between probability of detection and deployment of treatment in a realistic landscape. Our results show that there are optima that balance the expenditure on the identification of newly infected sites against the remaining resources available to treat the sites.

The conventional approach to dealing with emerging epidemics involves scaling up of the budget for control as the perception of the risk from the epidemic becomes more widely understood. This
was certainly the case during the Dutch elm disease epidemic in the United Kingdom in the 1970s (10) and also during the 50-y attempt to control white pine blister rust in the United States (13). Indeed, for SOD, the budget for control of the isolated epidemic in southwestern Oregon increased over time as the outbreak became larger (23). We show here, however, how front loading of the budget with greater expenditure in the earlier years is likely to be markedly more effective in bringing epidemics under control, increasing long-term cost-efficiency. However, in practice it might be difficult for a policymaker to justify a very large expenditure on an epidemic that currently remains small, particularly because estimates of future sizes are often subject to considerable uncertainty early in epidemics (39).

For SOD in CA, our analyses show—for the first time, to our knowledge—that statewide action to slow epidemic spread has not been feasible for some years. Management efforts to reduce impacts at local scales must now be the focus. However, we have shown how management starting 2002 could, in principle, have been successful. Our work illustrates how mathematical models can be used to optimize and assess the likely feasibility of management of established plant disease, particularly when there is a limited budget. It therefore complements a body of work on foot-and-mouth disease (25–27, 36), which also focuses on model-based optimization of large-scale management for an epidemic with long-distance dispersal. Particular challenges for SOD, however, include the broad host range of the pathogen, difficulties surrounding detection leading to very extensive cryptic infection, and the lack of any realistic control options other than host removal (e.g., no vaccination). The challenge—but also the opportunity—presented by increasing rates of introductions of plant pathogens (2, 40) and increasing acceptance of models by policymakers (31) is now to more routinely use the insights from models in the early stages of invading epidemics, when carefully optimized management can still make a difference.

Materials and Methods

Our stochastic epidemic model (Fig. 1A) tracks the density of PR-infected host across CA at 250 × 250-m spatial resolution, extending the model developed by Meentemeyer et al. (18) to include detection and treatment. The model (18) was parameterized using data on pathogen spread at local and state-wide scales and validated by predicting the infection status of positive and negative sites surveyed by the California Oak Mortality Task Force (41). The AUC (area under the receiver operating curve) value was 0.89, indicating very good performance. The model allows for spatial heterogeneity in host density, multiple host species with different susceptibility and infectivity (16), short- and long-range dispersal (42), and environmental drivers to affect transmission (19). Further details are in SI Appendix.

ACKNOWLEDGMENTS. We thank Richard Stutt, Stephen Parnell, and Matthew Castle for discussions; and Mark Calleja for administering a computer cluster. We acknowledge funding from the BBSC, DEFRA, NSF, USDA, and Gordon and Betty Moore Foundation.