The Magnitude of Regional-Scale Tree Mortality Caused by the Invasive Pathogen Phytophthora ramorum

Richard C. Cobb1, Sarah E. Haas2, Nicholas Kruskamp1,3, Whalen W. Dillon4, Tedmund J. Swiecki4, David M. Rizzo4, Susan J. Frankel7, and Ross K. Meentemeyer3,8

1Department of Natural Resources and Environmental Science, California Polytechnic State University, San Luis Obispo, CA, USA, 2Inland Fisheries Division, Texas Parks and Wildlife Department, Austin, TX, USA, 3Center for Geospatial Analytics, North Carolina State University at Raleigh, Raleigh, NC, USA, 4Agronomy Department, University of Florida, Gainesville, FL, USA, 5Phytosphere Research, Vacaville, CA, USA, 6Department of Plant Pathology, University of California, Davis, CA, USA, 7Pacific Southwest Research Station, Albany, CA, USA, 8Department of Forestry and Environmental Resources, North Carolina State University at Raleigh, Raleigh, NC, USA

Abstract
Forest pathogens are important drivers of tree mortality across the globe, but it is exceptionally challenging to gather and build unbiased quantitative models of their impacts. Here we harness the rare data set matching the spatial scale of pathogen invasion, host, and disease heterogeneity to estimate infection and mortality for the four most susceptible host species of Phytophthora ramorum, an invasive pathogen that drives the most important biological cause of tree mortality in a broad geographic region of coastal California and southwest Oregon. As of 2012, the most current field survey year, we estimate 17.5 (±4.6, 95% CI [confidence interval]) million tanoak (Notholithocarpus densiflorus) stems were pathogen killed with an additional 71 (±21.5) million infected. We estimated 9.0 million (±2.2) coast live oak (Quercus agrifolia) and 1.7 million (±0.5) California black oak (Quercus kelloggii) stems are disease impacted (mortality and infection combined). Lastly, our estimates suggest infection in 95.2 million (±8.6) California bay laurel (Umbellularia californica), which does not suffer mortality from infection and represents a critical source of continued spread. Prevalent infection as of 2012 suggests the cumulative number of disease-killed stems likely increased from 20.8 to 42.8 million between 2012 and 2019 for all species. While these impacts are substantial, most host populations occur in a yet to be invaded region of northern California indicating that the disease will intensify in the coming decades.

1. Introduction
Over the last century large-scale tree mortality from global change-associated drought, outbreaks of native insects and pathogens, and the introduction and spread of exotic, invasive insects and pathogens has caused unprecedented alterations to forest ecosystems around the globe (Allen et al., 2015; Seidl et al., 2017; Smilker-Williamson et al., 2019). Since the colonial period, North American forests have undergone repeated introductions of exotic pathogens (Garnas et al., 2011; Rizzo et al., 2002; Weed et al., 2013) notable for the significant mortality they have and continue to cause as well as the difficulty in forecasting their impacts (Loo, 2009; Lovett et al., 2006; Simler-Williamson et al., 2019). The economic, ecological, and social consequences of these nonnative forest pathogens present a tremendous challenge to a diverse range of societal interests including those of natural resource managers, insurance providers, and homeowners (Lovett et al., 2016; Pimentel et al., 2005).

Quantitative predictions of host infection and mortality are fundamental to mapping risk of disease emergence, forecasting impacts to ecosystem function, and warning of threats to human wellbeing emerging from deleterious interactions with fire or zoonotic diseases (Boyd et al., 2013; Lamsal et al., 2011; Metz et al., 2011). However, estimating these basic characteristics of tree mortality at broad spatial scales is extremely challenging for at least two reasons. First, geographic disequilibrium is inherent to biological invasions. This is the condition where invasion into suitable habitat and or host populations is ongoing and the current extent most strongly reflects the history of introduction as opposed to biological or epidemiological drivers. Second, the tree mortality events of greatest interest to managers are typically geographically expansive, making monitoring costly and fraught with technical and logistical constraints. These problems are often intertwined and make achieving unbiased estimates of incidence and distribution difficult (Holdenrieder...
et al., 2004; Meentemeyer et al., 2012; Thompson et al., 2016). Furthermore, accurate pathogen detection and temporal resolution of infection are often unattainable with remote sensing technology, particularly for generalist pathogens with a broad host range that frequently cause minor symptoms or those that resemble prevalent but innocuous endemic pathogens (Davidson et al., 2011; Hansen et al., 2008, 2012). Lastly, forecasts of landscape or regional tree mortality is often impeded by a lack of spatially explicit data on the abundance of tree populations. Although formidable, this list of challenges also provides a roadmap for overcoming them through data integration from diverse sources such as published data sets, long-term monitoring plot networks, publicly available geospatial data, and agency and citizen science data (Hartmann et al., 2018; Meentemeyer et al., 2012, 2015).

In this study, we leverage data from multiple sources to estimate infection and mortality caused by *Phytophthora ramorum*, an invasive forest pathogen spanning central coastal California to southwestern coastal Oregon. *Phytophthora ramorum* causes the forest disease sudden oak death and was first documented circa 1990 in the greater San Francisco Bay metropolitan area (Rizzo et al., 2002). In the decades since, sudden oak death has reached epidemic levels in coastal forests of California and Southwestern Oregon and currently represents the most important biological cause of tree mortality within the region in terms of area impacted, annual spread rate, and severity of stand-level mortality (Figure 1; Cunniffe et al., 2016; Lamsal et al., 2011; Metz et al., 2012). Currently, *P. ramorum* extends across approximately 800 km from a southern extent in coastal forests of the Big Sur ecoregion (California) to an isolated outbreak in Curry County Oregon (d) are an example of ongoing landscape-level management response to the disease.

![Figure 1. The impacts of sudden oak death in California and Oregon. High stand densities typical of many at-risk tanoak forests (a) in California and Oregon are poised for intense disease impacts frequently observed at the stand (b, Big Sur, photo K. Frangioso) and landscape levels (c, Sonoma County, Credit: U.S. Forest Service, Pacific Southwest Region, Forest Health Protection) in central California. In (c), healthy tanoak trees appear yellow-green while conifers appear dark green. Extensive and aggressive disease mitigation measures in Curry County Oregon (d) are an example of ongoing landscape-level management response to the disease.](image-url)
high quality region-scale *P. ramorum* distribution data sets in western U.S. ecosystems (Cunniffe et al., 2016; Hansen et al., 2008, 2019; Meentemeyer et al., 2015).

*Phytophthora ramorum* affects a broad host range of over 130 species of trees, shrubs, herbs, and ferns, many of which occur commonly throughout the currently invaded and at-risk area or are moved long distances via the nursery industry (Garbelotto & Hayden, 2012; Rizzo & Garbelotto, 2003). Worldwide, *P. ramorum* hosts fall into three general epidemiological classes, (1) hosts which transmit the pathogen but infections are not lethal, (2) hosts that transmit and infections are lethal, and (3) hosts where infection is not transmitted but is lethal. In the western United States these epidemiological classes are exemplified (respectively) by California bay laurel (*Umbellularia californica*), tanoak (*Notholithocarpus densiflorus*), and several red oak species particularly coast live oak and the closely related species Shreve’s oak (*Quercus agrifolia* and *Q. parvula* var. *shrevei*, respectively) as well as California black oak (*Q. kelloggii*). Bay laurel is the most susceptible and competent known sporulation-supporting species and represents substantial risk of establishment and spread in wildlands (Davidson et al., 2011). Tanoak is also critical to pathogen establishment risk and spread because it is both a competent sporulation-supporting species and is prevalent in much of northern coastal California (Figure 1; Cunniffe et al., 2016; Lamsal et al., 2011). Most significantly, infected tanoak stems suffer mortality typically within 2–8 years of infection, driving rapid and extensive mortality at the stand scale (Cobb, Filipe, et al., 2012; Metz et al., 2012). True oak species (*Quercus* sp.) survive longer with infection and are more likely to recover compared to tanoak, but mortality has nonetheless been extensive and severe in these forests (Brown & Allen-Diaz, 2009; Dodd et al., 2008; Metz et al., 2012). Pathogen caused mortality can approach 100% of canopy trees in some areas, resulting in increased fuel loads, increased fire severity, and reduced productivity (Figure 1; Cobb et al., 2013; Metz et al., 2011; Valachovic et al., 2011). While ecosystem changes have been documented throughout the current range of invasion, the overwhelming majority of at-risk forests have not yet been invaded (Cunniffe et al., 2016; Lamsal et al., 2011).

In the decades since mortality of tanoak and coast live oak were first documented, sudden oak death has reached epidemic levels throughout its range. Data sets derived from repeated stand censuses have identified factors influencing mortality and infection rates (Cobb, Filipe, et al., 2012; McPherson et al., 2010; Metz

**Figure 2.** Data acquisition, model building, and estimation. Field plot and spatial data were acquired across a broad region of invaded and uninvaded forest in coastal California (a, insets: i, Sonoma plot network; ii, Big Sur plot network) and evaluated in an information-theoretic framework of model selection and model averaging (b). Validated models were applied in a geographic information system (GIS) describing landscape epidemiological and environmental variation (c), and region-scale impacts were estimated by combining these model outputs with a previous model of host distributions (d). Each example map highlights a section of an ongoing and currently isolated outbreak in Curry County, Oregon.
et al., 2012). In combination, these longitudinal data sets have provided estimates of mortality across stands and ecological regions through time, but the regional extent of disease impacts has remained speculative and is a critical information shortcoming that hampers management. Our goal in this analysis was to develop spatially explicit estimates of infection and mortality for the four host tree species most greatly impacted by *P. ramorum* in California and Oregon, the largest and most destructive tree mortality event caused by this pathogen. To overcome potential bias due to geographic disequilibrium and data limitation, we integrated empirical data on infection and mortality from multiple plot networks, remotely derived geospatial data, publicly available pathogen distribution data sets, and previously published spatial estimates of host density (Cobb, Filipe, et al., 2012; Lamsal et al., 2011; Meentemeyer et al., 2008, 2015; Metz et al., 2011). We aimed to (i) determine the relative influence of environmental factors, community composition, and pathogen distribution on disease impacts for our four focal host species and (ii) understand if robust regional estimates of impacts could be derived from these data sources (Figure 2). We then apply these results to glean insights into management strategy and management effectiveness for this disease and develop guidance for models of other regional-scale tree mortality events.

## 2. Materials and Methods

### 2.1. Host and Pathogen Systems

Our focal coastal California and Southwest Oregon forests are classified as Mediterranean-type climatic conditions with cool, wet winters and warm, dry summers that exhibit substantial variability in average annual precipitation (17–458 cm) and average temperature (minimum = −5.7 to 12.7°C; maximum = 5.1 to 29.8°C). Host species are highly heterogeneous within the study area with variation in local occurrence, density, and age or size class structured by historical factors including logging and fire, and gradients of temperature and moisture (Bowcutt, 2011; Cobb, Filipe, et al., 2012; Davis et al., 2010). Host densities vary dramatically within relatively small areas (0.25–10 km) from 200 to 6,000 stems per hectare, particularly 10–30 years post-fire or post-clear cutting (Cobb et al., 2017; Metz et al., 2012). Each of our four focal hosts are relatively poorly studied due to a perceived lack of economic value; however, each species is a common canopy tree in the study region, can reach sizes exceeding 100 cm diameter at breast height (1.35 m, dbh [diameter at breast height]), and ages in excess of 100 years (Bowcutt, 2011; Cobb, Filipe, et al., 2012; Waring & O’Hara, 2008).

*P. ramorum* inoculum produced on host tissue is spread at several spatial scales and by different mechanisms. Local dispersal (meters) occurs primarily through detachment of sporangia—a structure containing multiple spores—with passive dispersal during rain events via rain splash and canopy throughfall (Davidson et al., 2005). Intermediate-distance dispersal of sporangia and zoospores has been shown to occur up to 2 km using neutral molecular markers and is thought to be associated with turbulence during storm events (Mascheretti et al., 2008). Lastly, rare long-distance dispersal (10–500 km) occurs primarily through the movement of infected plant material, often associated with the ornamental plant trade (Goss et al., 2009; Grünwald et al., 2012). Sporulation is most active during warm, wet springs, creating variation in pathogen spread, population sizes, and genetic variability among years (Davidson et al., 2005, 2008; Eyre et al., 2013).

### 2.2. Host and Disease Empirical Data

Our plot-level models are developed and fit exclusively with data collected in California (Figure 2). We integrated plot data from four long-term monitoring networks (*N* = 812 plots) established between 2000 and 2007 along coastal California from Monterey County to Del Norte County consisting of four distinct plot networks all established to gather data on disease dynamics over time (Figure 2). The North Coast and Big Sur plot networks consist of 318 and 264 circular 500 M2 plots, respectively while the Sonoma Mountain plot network encompasses 198 square 225 m2 plots. Finally, we used a subset of 35 plots (circular, 200 m2 each) from the Phytosphere Research network to obtain information on coast live oak and California black oak disease impacts in an otherwise under sampled portion the San Francisco Bay area. All plots included in our data set were constrained to 50 m minimum distance from their nearest neighboring plot, a detailed map of the plot networks can be found in the supporting information.

The distribution and number of plots restricts the plot survey frequency both within and among networks. The most recent sampling dates for the North Coast plots ranged from 2004–2012 with 147 plots (46%) last
sampled in 2004. In the Big Sur network, the most recent data are from a date range between 2006 and 2011, whereas all Sonoma plots were last sampled in 2011 and Phytosphere plots in 2012. Preliminary analyses suggested improved model performance when the oldest survey (2004) was omitted. As such, we only include data from 2005–2012 in our analyses ($N = 668$ plots), and use survey information only from the most recent visit dates. During plot visits, stems $\geq 1$ cm dbh were tagged, measured, identified to species, assessed for health status (live/dead), and $P. \text{ramorum}$ infection (symptomatic/nonsymptomatic). *Phytophthora ramorum*-like symptoms (leaf spots and bleeding cankers) are common host responses to a range of pathogens; therefore, we only considered these symptoms to be indicative of $P. \text{ramorum}$ when the pathogen had been previously isolated by direct culturing from symptomatic plant tissue. For plots with confirmed infections, infection and disease-caused mortality was recorded and tracked at the individual stem scale (see Cobb et al. [2012] and Metz et al. [2012] for more information). Most of the trees in these systems are multi-stemmed and $P. \text{ramorum}$ causes mortality at the stem scale, often killing up to 90% of above ground biomass without killing below ground structures (Figure 1; Cobb, Filipe, et al., 2012; Metz et al., 2012). A total of 10,346 stems were monitored across the four networks, in which 4,676 (45%) were tanoak (across 206 plots), 2,586 (25%) were California bay laurel (across 352 plots), 2,410 (23%) were coast live oak (across 307 plots), and 289 (3%) trees were California black oak (across 102 plots).

### 2.3. Spatial Environmental and Host Information

Six variables describing climate, topography, and $P. \text{ramorum}$ epidemiology were estimated on a 1 ha grid for California and Oregon, a spatial resolution which matches that of host density maps (Lamsal et al., 2011). We upscaled a 30 m USGS National Elevation Dataset (digital elevation model—DEM) to the 1 ha grid cell and calculated two topographically determined environmental variables: topographic moisture index (TMI) which characterizes potential soil moisture and solar radiation index (SRI) which characterizes solar insolation for the site (see Lamsal et al., 2011). Climate variation from 1990–2010 was described using rainy season (1 November to 31 May) total precipitation, and minimum and maximum temperature, as 30-year average annual values (1981–2010) downscaled to 1 ha resolution from the PRISM model (Daly et al., 2001; Flint & Flint, 2012). Finally, we included competent host density, the sum of tanoak and bay laurel densities, the principle sporulation-supporting species (stems ha$^{-1}$).

### 2.4. Pathogen Spatial Heterogeneity

Pathogen dynamics were estimated on two spatial scales, the first is simply the record of cells with known infections (invasion status) and the second, an estimate of local inoculum pressure known as force of infection (FOI) which aims to account for geographic disequilibrium and invasion uncertainty simultaneously. Combined, these parameters represent high-confidence pathogen distribution data (invasion status) combined with the likelihood of invasion in unsampled locations (FOI). Force of infection ($F_i$) was estimated for each hectare across the range of each host distribution with a negative exponential dispersal kernel:

$$F_i = \sum_{k=1}^{N} \exp \left( \frac{-d_{ik}}{a} \right)$$

where $d_{ik}$ is the Euclidean distance between each potential source of infection $k$ and target cell $i$, and $N$ is the total number of target cells across the study extent. The parameter $a$ ranges from 1–100 and modifies the form of the dispersal kernel to reflect high or low dispersal limitation; we found the strongest evidence for $a = 25$ which is consistent with previous studies (Havel et al., 2002; Meentemeyer et al., 2008). Geographically referenced data of all known invaded sites as of 2012 ($\sim 3,120$) were used in the final estimate of FOI and invasion status.

### 2.5. Hypothesis Testing and Impact Estimation

We constructed five models of disease impacts for the four host species which reflect their differing epidemiological roles or limitations of the data sets. Infection was assessed for bay laurel and tanoak with mortality also separately modeled for tanoak. Infection and mortality could not be separately estimated for our focal oak species because both impacts occur at much lower frequencies compared to bay laurel and tanoak. Therefore, we estimated “disease-impacted” oak stems, a category combining infected and dead stems. Within each model the respective outcome variable, $y(s)$, is binomial proportion of the total number of trees.
for each focal species, \(n(s)\) per plot. These binomial proportions where then fit to generalized linear models (GLM) with a binomial distribution (logit link function) and quasibinomial error distribution.

We implemented a distance-based rule for model development to additionally reduce potential bias in parameter estimates emerging from pathogen geographic disequilibrium: all estimation was constrained to no more than 4 km away from a known confirmed \(P. \text{ramorum}\) infection. This rule is based on previous sudden oak death research showing that the majority of rare long-distance pathogen dispersal events occur within 2 km, which limits projections to a spatial scale of local invasion documented in multiple studies (Davidson et al., 2005; Hansen et al., 2008; Mascheretti et al., 2008). The decision also led to exclusion of an additional 145 plots (18%) primarily located in the three northernmost California counties (Mendocino, Humboldt, and Del Norte) where the pathogen had not yet colonized as of 2012.

For all models, we excluded predictor variables with a correlation coefficient \(|r| > 0.60\) to minimize potential multicollinearity. This occurred in only one case, rainy season average maximum temperature and precipitation \((r = -0.71)\). This problem was overcome by only including precipitation given that field and laboratory experiments have clearly demonstrated its importance to pathogen spread (Davidson et al., 2005, 2008). We used the R package “MuMIn” (Barton, 2013) to generate a set of models with all possible combinations of the terms including host, environment, and pathogen (prevalence and local inoculum). All models were identical except for Bay laurel where exclusion of FOI improved model fit. Parameter estimates from models with similar support (\(\Delta QAIC < 2\)) were averaged, a method to avoid selecting a spurious “best” model from a set with similar relative likelihood (Burnham & Anderson, 2004). We applied model-averaged values to estimate the probability of the respective disease outcome for each hectare across the prediction domain using the R package “raster” (Hijmans, 2014). Each probability raster was then multiplied by host density estimates from Lamsal et al. (2011) to produce a map of the number of disease-impacted trees per hectare across California and Oregon (constrained to a 4 km buffer of confirmed \(P. \text{ramorum}\) infections; Figure 2). Finally, estimates of disease impacts were summarized at the range-wide, state, and ecoregion extents. For the final averaged models, we derived 95% confidence intervals of each prediction by creating explicit models of prediction error across a continuous range of disease impacts as opposed to using model weights to attain the unconditional variance estimator (see Burnham & Anderson, 2004). We found linear changes in prediction precision for all models and calculated prediction error based on 95% confidence intervals for each individual cell.

### 2.6. Model Performance and Validation

For each species, adequacy of model structure was evaluated on the basis of model overdispersion using goodness of fit statistic for model selection procedures \(\hat{c}\) — the ratio of the global model chi-square and degrees of freedom. Overdispersion, often referred to as variance inflation, is the condition where observed variance is greater than that of the theoretical model being to fit the data. Overdispersion can be a serious challenge to inference and prediction as it can introduce significant bias to both modeling goals. In our analysis, models with \(\hat{c} > 6\) were rejected as structurally inadequate and untrustworthy. Models suitable for inference and prediction were those with \(\hat{c} < 4\) while those with \(4 < \hat{c} < 6\) suggest structural improvement is possible but are acceptable for preliminary estimates (see Burnham & Anderson, 2004). We also applied two diagnostic metrics on the model residuals: (i) the root mean squared error (RMSE) as a measure of overall prediction quality based on differences between true and predicted values and (ii) normal distribution of error and constant variance. We also estimated the precision of model predictions by calculating Pearson’s \(r\) for predicted and observed disease-impacted trees in each plot. Inference on the model parameters was based on their respective effect sizes \((\lambda_i)\): the ratio of the estimated coefficient and the respective adjusted standard error (Burnham & Anderson, 2004). We extracted Akaike weights \((\omega_i)\) for each model parameter and highlight those where \(\lambda_i \geq 1\) and \(\omega_i = 1\) which indicate that the parameter has a relatively large effect on the outcome and it was present in every candidate best model \((\lambda_i\) and \(\omega_i\), respectively). Variables with \(\lambda_i < 1\) but \(\omega_i = 1\) and \(\lambda_i \geq 1\) but \(\omega_i < 1\) are also discussed although their importance is less clear cut because their effects are relatively weak or they do not appear in every model included in the averaged parameters, respectively. A limitation of both the model selection and residual-based performance statistics is that the same data is being used to both fit and test the model, which may result in optimistic performance estimates. To address this, we conducted a \(k\)-fold cross validation by randomly selecting 10% of the data as a hold-out validation.
data set, fitting and subsequently assessing model fit ($\hat{c}$) and measuring prediction precision of the held-out data (Pearson’s $r$; Arlot, 2010). The cross validation was bootstrapped 100 times with independent and randomized $k$-fold selections for each training and testing data set ($k = 10$ each validation, 1,000 total model fits). This approach produces a distribution of model fits for selection ($\hat{c}$) and prediction accuracies ($r$) which provides insight into the degree that specific models or predictions are robust to eccentricities of the data sets.

3. Results

3.1. Model Fit and Validation

Our model selection and estimation procedures produced a range of model fit and prediction accuracy estimates (Figure 3). Within the $k$-fold model evaluation, overdispersion estimates ($\hat{c}$) had a relatively low range of variation, particularly when $\hat{c} < 3$. We found little evidence of systemic structural problems for all models (e.g., $\hat{c} < 4$) except for the model of tanoak infection for which $\hat{c} = 5.4$ and where some of the $k$-fold test models had $\hat{c}$ values very close to our threshold for rejection but was never exceeded (Figure 3). Further exploration of alternative model structures, underlying model distributions, or alternative predictors did not reveal any clear opportunity for model improvement. While this relatively high $\hat{c}$ value suggests modeling of tanoak infection processes at broad spatial scales could be improved, we include predictions from the model as a preliminary estimate.

Precision for predicted values ($r$) ranged broadly across test models and was highest for bay laurel and lowest for both oak species (Figure 3).

3.2. Cumulative Sudden Oak Death Impact Estimates as of 2012

Cumulative infection in California bay laurel was estimated at 95.23 million stems (±8.59, 95% CI [confidence interval]), which represents ~21.7% of the total population (Table 1). For tanoak, 71 million stems were infected (±21.5, 95% CI) while another 17.5 million stems were estimated to have been killed (±4.6, 95% CI); these represent ~4% and 1% of the total tanoak population, respectively. Cumulative disease impacts in coast live oak (dead and infected) was estimated to be 9 million stems (±2.2, 95% CI) and 1.7 million California black oak (±0.5, 95% CI), representing 2% and 0.1% of the respective total populations. For these Quercus species, 69% of trees from both species had died following infection with the remaining 31% infected but living as of the most recent survey.

3.3. Spatial Patterns of Disease Impacts

Tanoak is notable for several geographic areas with distinctly intense disease including an isolated outbreak in Curry County Oregon, the Big Sur region, the Santa Cruz Mountains, and the collective area of Sonoma and Marin counties (Figure 4). Tanoak mortality closely corresponds to the distribution and intensity of infection (Figure 5). Disease hotspots for tanoak represent substantial infection when viewed at the state or ecoregion scale, particularly when compared with estimated population sizes reported by Lamsal et al. (2011). The most intense infection in tanoak occurs in the California Chaparral and Oak woodland ecoregion with 13.7% of the total population infected, followed by the Coast Range ecoregion where 4.8% is infected (Table 1). In the Klamath ecoregion infection occurs in an estimated 2.9% of tanoak stems which reflects the overall lower prevalence of infection as well as the ecoregion’s extensive uninvaded forests (Table 1). The Klamath and Coast Range ecoregions span the California and Oregon border thus estimates include both states. Estimated infection in tanoak was substantial in Oregon regardless of the lens through which this is viewed. Within the model estimates, Oregon forests represent an estimated 41.3% of the total infections in tanoak, or 38.8% of the total infected and killed tanoak stems combined. In relative terms,
7.9% of Oregon tanoak were estimated to be infected compared to 2.9% of the California tanoak stems. These patterns demonstrate the extent to which invasion is limited in California relative to the host population, despite the substantial impact overall. Oregon has also aggressively worked to eradicate *P. ramorum* because our models do not account for these actions, the estimates represent expected disease in absence of control measures.

Estimated infection in bay laurel is primarily restricted to California which also encompasses the vast majority of the total population (~96%). Notably high within-stand infection levels occur in Sonoma County and the Big Sur region where 100% infection prevalence was often recorded within our study plots (Figure 4). Disease in coast live oak is also notably high in the Sonoma and Big Sur regions, while impacts to California black oak were predicted primarily in Sonoma County and along the Sonoma County border with Napa County (Figure 6). Disease impacts to our focal oak species rapidly declines north of Sonoma County as

Table 1
Cumulative Sudden Oak Death Impacted Stems for the Four Major Host Species in California and Oregon Forests as of 2012 (in Millions of Stems)

| Geographic extent | Bay laurel infection | Tanoak infection | Tanoak mortality | Coast live oak | California black oak |
|-------------------|----------------------|------------------|------------------|----------------|---------------------|
| Range-wide        | 95.23 (8.59)         | 71.01 (21.50)    | 17.51 (4.60)     | 9.02 (2.16)    | 1.75 (0.51)         |
| **Estimation by state** |                   |                  |                  |                |                     |
| California        | 91.82 (8.28)         | 40.96 (12.43)    | 13.19 (3.47)     | 9.02 (2.16)    | 1.75 (0.51)         |
| Oregon            | 3.4 (0.31)           | 30.05 (9.07)     | 4.32 (1.13)      | NA             | 0.00 (0.00)         |
| **Estimation by ecoregion** |               |                  |                  |                |                     |
| Coast range       | 24.59 (22.22)        | 41.12 (12.45)    | 11.44 (3.00)     | 1.74 (0.41)    | 0.33 (0.10)         |
| Klamath Mountains | 1.08 (0.1)           | 21.94 (6.62)     | 2.91 (0.76)      | 0.01 (0.00)    | 0.01 (0.00)         |
| CA Chap. and oak woodland\(^a\) | 69.53 (6.27) | 7.95 (2.43)      | 3.16 (0.84)      | 7.27 (1.74)    | 1.41 (0.41)         |

Note. Estimates are categorized by range-wide, state, and ecoregions for California and include 95% confidence intervals in parentheses. NA = the species does not occur in natural ecosystems within this region.

\(^a\)California Chaparral and Oak Woodland Ecoregion.

Figure 4. Estimated densities of tanoak (left) and California bay laurel (right) stems infected by *Phytophthora ramorum* as of 2012. Note the different range in density for each species.
coast live oak reaches its species distribution limit in this region and the pathogen has not invaded California black oak forests beyond the southern half of Mendocino County.

### 3.4. Environmental and Host Disease Drivers

Our disease models suggest strong support for several epidemiological and environmental drivers of disease as well as several common inferences regarding environmental drivers of infection prevalence in tanoak and bay laurel specifically (Table 2). For both species we found strong support for positive associations of infection with competent host density and wet-season average minimum temperature. In addition, infection in bay laurel was positively associated with average total precipitation while topographic moisture variability (TMI) did not significantly affect infection in either species (Table 2). For bay laurel, we found strong support for a positive association between pathogen prevalence and SRI while infection in tanoak appears to be negatively associated with this parameter. These environmental factors describe sites with greater precipitation and warmer conditions during the period of transmission (January–May) which favors infection in bay laurel and tanoak. However, the opposing relationships for these two species regarding SRI suggests temperature conditions over the course of the full year further modify infection, possibly with higher temperatures reducing over-summer pathogen survival for tanoak but increasing it in bay laurel.

Plot-level drivers of tanoak mortality were notably different from infection in tanoak and bay laurel. In comparison, we found little support for strong relationships between tanoak mortality, competent host density, precipitation or temperature. However, we found support for a positive relationship of mortality with...
topographic differences in moisture availability (TMI) and a negative relationship for solar insolation (SRI). We also found support for a positive relationship between tanoak mortality and invasion status. These patterns suggest invasion of wetter sites leads to greater mortality but that interception of solar radiation and possibly temperature outside of the seasonal transmission period may mitigate some of these impacts. Mortality in tanoak is driven by pathogen occurrence in these models; therefore, infection and mortality are closely associated at broad spatial scales (Figures 4 and 5). However, variation in the importance of environmental drivers between infection and mortality suggests some edaphic conditions may result in longer than expected tanoak survival times following infection.

In disease prevalence models for coast live oak, we found strong support for positive relationships of disease impacts with competent host density and FOI but no support for the effects of any of our environmental variables (Table 2). For disease prevalence in California black oak we found some support for positive

![Figure 6. Estimated density of disease-impacted (infected and pathogen killed) coast live oak (left) and California black oak stems (right). Note the difference in range of density for each species.](image)

### Table 2

| Species          | Disease measure      | Invasion status | Host density | TMI | SRI  | RS average precipitation | RS average T minimum | Force of infection |
|------------------|----------------------|----------------|--------------|-----|------|--------------------------|----------------------|-------------------|
| Tanoak           | Infection            | 0.02 (1)       | 3.56 (1)     | −0.69 (0.47) | −2.71 (1) | 0.4 (0.3) | 3.29 (1) | 0.58 (0.4) |
| Tanoak           | Mortality            | 5.14 (1)       | −0.27 (0.19) | 2.82 (1)     | −3.0 (1)  | −0.9 (0.15) | 0.02 (0.14) | 0.02 (0.14) |
| Bay laurel       | Infection            | 0.02 (1)       | 7.33 (1)     | −0.7 (0.5)   | 4.34 (1)  | 2.87 (1)  | 4.72 (1)  | NA               |
| Coast live oak   | Infection and mortality | 0.55 (0.34) | 2.48 (1) | −0.07 (0.09) | −0.01 (0.09) | −0.18 (0.1) | 0.31 (0.23) | 3.42 (1) |
| Black oak        | Infection and mortality | 1.1 (0.71)  | 0.55 (0.36) | 0.1 (0.05)  | 0.08 (0.05) | −0.57 (0.36) | 1.24 (0.76) | 0.08 (0.1) |

Note: Parameter values with $\lambda_i > 1$ and $\omega_i = 1$ are highlighted in bold, while those with lower support ($\lambda_i > 1$ and $\omega_i < 1$) are italicized. Parameter name abbreviations: TMI = topographic moisture index; SRI = solar radiation intensity; RS = rainy season; T = temperature ($^\circ$C).
associations with average minimum temperature and invasion status in that effect sizes were $> 1$ but factors were not included in every model (e.g., $\omega < 1$). However, black oak had the greatest number of total models with similar support (14) which suggests these factors, although not clear-cut, should not be discounted. We found no support for the effect of any other environmental variables on disease prevalence in black oak.

Overall, these models point to variable levels of support for specific epidemiological or environmental variables across these four host species. That coast live oak disease impacts are positively associated with competent host density is unsurprising given that neither this species nor California black oak are known to support sporulation and can only acquire infection from neighboring trees. Thus, while disease impacts in tanoak and bay laurel are broadly prevalent in invaded areas, smaller-scale variation in inoculum appears to have a greater influence on infection in oaks. This difference likely reflects the competency for bay laurel and tanoak to transmit the pathogen and thus increase inoculum on isolated trees even during rare long-distance dispersal events where as mortality in tanoak may be more strongly associated with local inoculum pressure as suggested by the strong support for positive effects of invasion status.

4. Discussion

4.1. How Destructive Is Sudden Oak Death and How Much Worse Can the Disease Become?

For *P. ramorum* our models suggest 166.2 million total stems were infected by this pathogen as of 2012. Of these, 71 million stems were tanoak for which mortality rates of infected stems average $\sim 5\%$ on decadal scales (Cobb et al., 2010). This suggests that as of 2019 at least 21.4 million of these infected stems have since died which would result in 38.94 million tanoak stems killed since the outbreak emerged or $\sim 2.2\%$ of the total population. For both oak species, dead stems accounted for 69% of the total impacted, or approximately 6.2 and 1.2 million coast live oak and black oak, respectively killed by 2012. Given mortality rates of infected coast live oak and black oak have been measured at $\sim 3\%$ and $\sim 5\%$ per year, respectively, this suggests an additional 1.12 million coast live oak and 0.16 million California black oak have died since 2012 (Brown & Allen-Diaz, 2009; McPherson et al., 2010). In total, it is likely that at least 47.62 million individual stems have been killed across all species. However, in light of the substantial population of infected stems and that pathogen spread is an ongoing process both at the stand and landscape levels this estimate should be considered a lower bound.

Both our direct estimates of cumulative impacts as of 2012 and the projected impacts as of 2019 are alarmingly high. These estimates should be viewed through the lens of extensive stand and ecosystem studies which frequently demonstrate transformed fuel profiles, reduced stand biomass, and altered species composition (Brown & Allen-Diaz, 2009; Cobb, Chan, et al., 2012; Metz et al., 2012). Furthermore, annual surveys within subsets of our data set and stand-to-regional epidemiological modeling have documented increasing rates of infection showing this disease is progressively worsening over time (Cobb et al., 2010, 2012; Cunniffe et al., 2016; Meentemeyer et al., 2011). The cumulative estimate of 166 million infected tanoak and bay laurel almost certainly has increased since 2012 as these infections represent tremendous capacity for regional spread and local intensification. About 1.7 billion total tanoak stems alone are at risk across the entire study area and represent approximately 68 Tg living-biomass carbon, for which a substantial proportion is likely to be released to the atmosphere by the disease (Fei et al., 2019; Kurz et al., 2008; Lamsal et al., 2011). Much of this expansive at-risk forest is in northern California and Southwestern Oregon, the portion of our study area where pathogen invasion is relatively limited but where the potential for devastating landscape-level mortality is greatest.

4.2. Insights Into Other Tree Mortality Estimates

Given the extensive data on pathogen infection prevalence and epidemiology, field quantification of mortality dynamics, and detailed models of host distribution, extending our approach to other tree mortality events is a nontrivial task but this study demonstrates the possibility and scope. Our estimates rely on a geographically extensive common-methodology plot network monitored over a relatively long period (2004–2012). This surveillance network was particularly critical in this case because sudden oak death impacts many areas that are poorly represented by publicly available forest health and forest structure data sets (e.g., USDA Forest Service Forest Inventory and Analysis—FIA; Cobb, Filipe, et al., 2012; Haas et al., 2016; Lamsal et al., 2011; Metz et al., 2012). Mustering data sets which encompass the full range of emerging mortality drivers is critical to developing similar models for other mortality events. Gathering essential empirical
Many biological characteristics of tree mortality events will complicate reliable forecasting. Stymied management as a result of incomplete detection is a critical problem for plant, animal, and human disease including *P. ramorum* in the nursery industry, Leishmaniasis, COVID-19, and many others. This critical shortcoming results from failures to detect cryptic (latent) infection and/or survey efforts that do not match the scale of invasion (Filipe et al., 2012). We have little doubt our choice to classify infections based on laboratory confirmation of the pathogen results in underestimation of overall infection. In addition, our surveys do not account for infection in trees which died before plot establishment, infection and death between plot surveys, and false-negative diagnosis. Incomplete detection is particularly problematic for invasive species and probably results in an underestimation of mortality for many invasive insect and disease-driven mortality events (Fei et al., 2019). In our data set, underestimation is most likely in locations with high inoculum levels and where surveys occurred on intervals greater than 5 years including the Sonoma and Santa Cruz County regions. Previous studies have also found evidence of infection underestimation for both oak and tanoak forest types despite their epidemiological differences (Brown & Allen-Diaz, 2009; Cobb, Filipe, et al., 2012).

Tree mortality estimation must also be approached as an adaptive exercise. While many natural resource scientists may find this self-evident, overdispersion in our model for tanoak infection is a useful example of the importance of constant work to improve models as well as the inherent challenges. We could not address model overdispersion with readily available data (Figure 3) and this shortcoming represents the potential for improvement of model structure and predictions (Burnham & Anderson, 2004). Using two measures of pathogen establishment, both invasion status and inoculum loads implied by the FOI variable helped reduce overdispersion in all models except for bay laurel where model performance was improved by omission of FOI (Table 2). Bay laurel is the most susceptible of our host species and does not suffer mortality following infection which is likely to reduce detection bias. Furthermore, several independent studies have shown that cumulative infected tanoak and oak increases over consecutive surveys while infection in bay laurel is more likely to have already reached maximum detected levels during the first survey (Cobb, Filipe, et al., 2012; Haas et al., 2016; Metz et al., 2012). In all models except for infection in tanoak, epidemiological process is especially fast (infection in bay laurel) or relatively slow (disease in oaks, mortality in tanoak). These patterns imply our snapshot approach may be more accurate for particularly rapid or slow processes while estimation for intermediate rate processes may be more sensitive to the temporal nonlinearities characteristic of infectious disease. The challenge of documenting patterns and underlying drivers for emerging mortality events is likely to scale with the extent of the event and, like pathogen establishment and landscape inoculum levels, is likely to change over time (Cobb et al., 2017; Havel et al., 2002; Meentemeyer et al., 2008).

### 4.3. Management Implications

Models of *P. ramorum* establishment risk have unanimously suggested heightened infection probability along coastal forest ecosystems from southern central California northwards to isolated outbreaks in southwestern Oregon. These areas are the most conducive to outbreak in terms of epidemiological and environmental variables, although pathogen prevalence in these landscapes is relatively low. The most severe impacts have consistently been associated with time since invasion, for example Santa Cruz and Marin Counties as well as the Big Sur region (Figures 4–6; Mascheretti et al., 2008; Meentemeyer et al., 2011; Rizzo et al., 2005). Later occurring introductions in Humboldt County and Curry County, Oregon remain relatively isolated although both of these introductions have been met with major efforts to control pathogen spread.

In Oregon an aggressive, well-funded and well-organized campaign applied eradication-style treatments on 2,170 ha of land as of 2012. The campaign has continued through 2019 and combines systematic aerial and ground surveys which are followed up with eradication treatments to cull all hosts in an area extending at least 100 m beyond symptomatic plants (Hansen et al., 2019; Peterson et al., 2014). These treatment boundaries encompassed ~5 million tanoak stems as of 2012. Our models were fit exclusively with data from California while disease impacts in Oregon are estimated using local vegetation, environmental, and...
pathogen data. Some epidemiological differences between California and Oregon do occur, in particular California bay laurel does not appear to be as epidemiologically important in Oregon and is far less abundant (Hansen et al., 2008; Peterson et al., 2014). Of greater importance, estimates of tanoak infection represent disease impacts in Oregon without management and imply that the efforts have prevented infection in as many as 25 million tanoak stems—the total estimated for Oregon less the stems within treatment boundaries as of 2012. While continued spread of the pathogen has necessitated abandonment of eradication efforts in heavily invaded areas, infection levels remain low in actively managed areas (Hansen et al., 2019). Potential impacts to tanoak in southwestern Oregon, and Del Norte, and Humboldt Counties in California are very high due to the convergence of conducive environmental and host conditions (Table 2). However, our models suggest intervention has substantially slowed intensification of P. ramorum even though new isolated infections are detected every year (Hansen et al., 2008, 2019; Peterson et al., 2014). This suggests continued effort to contain pathogen spread is justified in Oregon and that similar campaigns could be effective in northern California as well.

Given previous research showing disease impacts can include changes to fuels, fire-caused mortality, and ecosystem productivity (Cobb, Chan, et al., 2012; Cobb et al., 2013; Metz et al., 2013), it is natural to ask how our estimates of cumulative mortality reported here translate to broader ecosystem health across the landscape as well as the implication for broader public policy. A recent analysis of invasive pest and pathogen impacts for the United States estimated as much as 76.7 Tg live host carbon occurs in counties currently invaded by P. ramorum (Fei et al., 2019). Assuming our estimate of 2.2% cumulative mortality scales with carbon, we estimate ~1.68 Tg C have been killed by the disease or about 0.084 Tg C per year since 1999 (all species; Lamsal et al., 2011). In a recent meta-analysis of forest carbon loss due to invasive forest insect and pathogen outbreak using USDA FIA data, Fei et al. (2019) did not find evidence of net annual loss for sudden oak death. As we previously noted, the Forest Inventory Analysis program has poor coverage in the regions with the greatest current P. ramorum impact (Lamsal et al., 2011). The FIA data also designate mortality on a whole tree basis as opposed to the stem-by-stem mortality caused by this pathogen. Sudden oak death impacts are thus almost certainly greater than those implied by the FIA data alone implying that changes in fuels, fire, and ecosystem processes documented at the stand scale are likely occurring throughout broad parts of the currently invaded range. While the current annual loss of forest carbon may not reach the thresholds for state-level action, transformations of structure that influence fire and other forest resources, and the potential for exponential infection increase likely justifies focused and sustained intervention at regional scales.

The estimate of sudden oak death we report here rivals recent tree mortality events in the western United States such as those associated with drought and insect outbreak (Preisler et al., 2017), but we stress again that these are a fraction of the potential impacts in the coming decades (Cunniffe et al., 2016; Meentemeyer et al., 2011). Our study represents a notable advance for responding to sudden oak death, yet the pathogen has continued to spread and kill trees in the years since the most recent measurements used in our models (2012). The overall impact of tree mortality on sequestered carbon, biodiversity, and other management issues will only be clear in hindsight for most tree mortality events. However, previous work has shown that proactive treatments are relatively inexpensive compared to costly fuel removals and ecosystem restoration necessitated by intense sudden oak death mortality (Cobb, Hartsough, et al., 2017). In the case of P. ramorum there is a clear risk associated with inaction (Table 1) and an equally clear opportunity to intervene against the dramatic impacts to Pacific Coast forests this pathogen is otherwise certain to cause in the coming decades.

Data Availability Statement

Data employed to develop predictive models, estimate impacts, and derived data products (host impact maps), code, and metadata for this study are fully available in the Dryad data repository (https://doi.org/10.5061/dryad.n8pk0p2rs).

References

Allen, C. D., Breshears, D. D., & McDowell, N. G. (2015). On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene. Ecosphere, 6(8), art129. https://doi.org/10.1890/ES15-00203.1

Arlot, S. (2010). A survey of cross-validation procedures for model selection. Statistics Surveys, 4, 40–79. https://doi.org/10.1214/09-ss054
