Environment and host infection history jointly predict disease risk in a multi-pathogen system

Scott, C. B.[1,*], Cleary, S.[1], Grunberg, R.[2], Halliday, F.[3], Joyner, B[4]., O'Keeffe, K.[5], Stiver, I.[3] & Mitchell, C. E.[1]

[1]Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[2]Wilson Center for Science and Justice at Duke Law, Durham, NC, USA

[3]Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR, USA

[4]Department of Parks, Recreation and Tourism Management, North Carolina State University, Raleigh, NC, USA

[5]Department of Biological Sciences, Lehigh University, Bethlehem, PA, USA

*corresponding author; cabscott@unc.edu

Abstract

Forecasting disease epidemics may require considering both abiotic and biotic conditions. Abiotic conditions can influence pathogen dispersal, survival, and infection, while prior infection by one pathogen species can alter host susceptibility to subsequent pathogens, creating historical contingencies. Yet, the relative importance of environmental conditions, within-host pathogen interactions, and their potential interplay in predicting seasonal disease dynamics remain underexplored. To better understand these interactions and improve ecological forecasts of disease risk, we analyzed more than 41,000 plant-level observations of three foliar fungal diseases on the grass species tall fescue in North Carolina, collected from 2017–2024. We built temporally explicit random forest models which for two of the three focal diseases accurately (>80%) predicted future disease dynamics from a complex abiotic and biotic predictor space. From these models, we identified key environmental thresholds that intensified annual epidemics. We then supplemented those machine-learning models with Bayesian hierarchical

models and survival analyses, finding that pathogen-pathogen interactions can be as important as environmental conditions in predicting disease risk. Furthermore, for two of the three focal diseases, prior infection by a different pathogen facilitated the subsequent infection and the strength of this facilitation was modulated by environmental context. From this study we draw two major conclusions, one ecological and one methodological. First, knowledge of a host's current disease state may be as important as local environmental conditions for predicting disease dynamics. Second, by integrating complementary modeling approaches, we can develop both predictive forecasts and mechanistic insight into the biotic and abiotic drivers of infectious disease.

Open Research Statement: All data and code used to generate the results contained in this paper are available on GitHub (https://github.com/cb-scott/WidenerLongTermAnalysis). Upon publication the final version of all scripts and datafiles will be archived on Zenodo with a permanent DOI.

Introduction

The success of pathogen transmission is jointly shaped by both abiotic environmental factors and the biotic context. During transmission, many pathogens are exposed to environmental conditions, such as temperature, moisture, or wind, that critically influence their ability to survive, disperse, or infect (Altizer et al., 2006). Upon encountering a host, a pathogen interacts with an established microbial community, potentially including pathogens that have already infected the host. These species interactions may affect the pathogen's infection success, including through within-host priority effects. Such priority effects can either decrease pathogen infection success via antagonistic interactions, increase infection success via facilitative interactions, or have no effect on infection success, making transmission outcomes potentially subject to historical contingency (Clay et al., 2020; Fukami, 2015). Historical contingency intersects with environmental factors to influence pathogen transmission, where the strength and direction of pathogen species interactions may vary with shifting environmental contexts (Vannette & Fukami, 2014). Together, these mechanisms create a temporally dynamic interplay between environment and community history that may shape pathogen epidemic trajectories (Hassell et al., 2021).

The relative contributions of environmental factors and historical contingencies to pathogen dynamics, as well as the importance of interactions between them, remain poorly resolved (Halliday et al., 2021). Separating these influences requires quantifying both the environmental niches of individual pathogen species and the strength of their interactions, yet these parameters are rarely measured together. Many pathogens occupy broadly overlapping abiotic niches, making it difficult to infer the role of historical contingency from co-occurrence data alone (Suh et al., 2024). Consequently, reliance on single-timepoint or solely correlative approaches often

conflates abiotic filtering with historical contingency, limiting our ability to identify when order of arrival genuinely shapes infection outcomes.

Disentangling these processes requires analytical methods that can accommodate collinearity among environmental predictors and non-monotonic relationships between those predictors and disease outcomes. Addressing both goals, machine learning algorithms capture complex, non-linear dynamic and offer robust predictions of future community states (Prasad et al., 2006). Here, we focus specifically on random forest algorithms, which are relatively simple to implement and offer insights from high-dimensional predictor data. By constructing many decision trees from random subsets of the data (hence, the name “random forest”), random forests not only identify the strongest predictors of a response but also reveal the environmental or biological thresholds at which these predictors alter model predictive performance. While random forests have recently been applied to predict future disease outbreaks, these models have focused solely on abiotic drivers of disease (Alves et al., 2025; Shah et al., 2023).

However, despite the power of random forest algorithms, the goal of these models is predictive accuracy rather than mechanistic understanding. Although methods exist to interpret model results (for example, see, Molnar, 2025), post hoc analyses are often required to identify important variable thresholds, and visualizing or interpreting interaction terms can be unintuitive. Often, experiments have additional design consideration which would be poorly captured by machine learning methods, such as the repeated sampling of locations or individuals through time. Supplementing machine learning with traditional statistical methods can enhance interpretability and predictive performance in plant biology, particularly when hierarchical structure, known biological constraints, or uncertainty quantification are important (Kick & Washburn, 2023). For example, Bayesian hierarchical models have been used to compare the

spatial and temporal drivers of disease and quantify uncertainty in unobserved disease states (Farnsworth et al., 2006; Gibson et al., 2004). Similarly, survival analyses offer a natural framework for time-to-event data in plant pathology, such as modeling incubation period length and disease symptom timing while accommodating censored observations (Scherm & Ojiambo, 2004).

Here, we first apply random forest algorithms to climatic and biotic variables to predict seasonal epidemic trajectories of three plant diseases in an old field over seven years. Then, we leverage Bayesian hierarchical models and survival analyses to fully utilize the structure of our data to gain a mechanistic understanding into the environmental and biotic drivers of disease. Through our combined modeling approach, we find that within-host pathogen interactions can be as important as environmental context in driving disease risk and identify key environmental thresholds affecting disease risk.

Methods

Study System

Tall fescue (*Lolium arundinaceum*) is a widespread cool-season grass that is grown for turf, fodder and forage, and is an ecologically dominant species in many unmanaged plant communities across much of the eastern United States (Natural Resources Conservation Service, 2001). We focus on three foliar fungal diseases that commonly coinfect tall fescue in the Piedmont region of North Carolina: anthracnose, brown patch, and crown rust. Anthracnose at our field site is caused chiefly, but not exclusively, by *Colletotrichum cereale*, which spreads

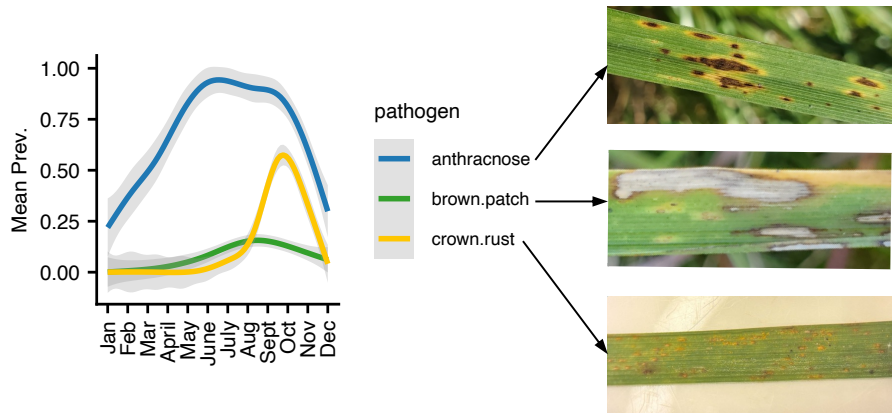

Figure 1. Mean yearly dynamics for anthracnose, brown patch, and crown rust diseases. For each disease, trendline gives cubic spline and 95% confidence interval through observations of disease prevalence in full dataset from 2017 to 2024. Anthracnose peaks earliest, in June, followed by brown patch in August, and crown rust in October. Photos at right give an example of each lesion type on tall fescue. *Photos: Susan Cleary, Dalia Chen*

primarily through rain-splash dispersal of asexual conidia (Crouch & Beirn, 2009). Brown patch is caused by *Rhizoctonia solani*, which persists in the soil, and disseminates through hyphal fragments and growth in soil and water films as well as sclerotia (Ogoshi, 1987). Crown rust is caused by *Puccinia coronata* and is dispersed via wind-borne spores. Together, these diseases affect foliar health and performance of tall fescue across diverse environmental conditions (Butler & Kerns, 2019a, 2019b, 2019c). All three diseases are visually distinct from each other in the field (Figure 1).

The three focal diseases exhibit distinct seasonal cycles. Anthracnose prevalence peaks in May, followed by brown patch prevalence peaking in August, and crown rust prevalence peaking in late October (Figure 1). These temporally staggered outbreaks create opportunities for coinfection and potential pathogen interactions that could influence host performance and disease severity over the growing season. Interactions between these pathogens in tall fescue can alter disease progression (O'Keeffe et al., 2021; Green, Grunberg, & Mitchell, 2024) and

influence host susceptibility to later-arriving diseases (Halliday et al., 2017). Building on this, we seek to quantify how the influence of pathogen–pathogen interactions have compared with abiotic context in shaping epidemic dynamics over 7 years.

Disease Survey Data

All disease surveys were conducted between 2017 and 2024 at Widener Field in Duke Forest, North Carolina. Data from multiple past experimental field studies were combined for the analysis (Grunberg et al., 2025). Some past experiments involved the treatment of experimental field plots with fungicide. For these datasets, only control (i.e. no-fungicide) plots were retained for analysis. In some years, disease presence was measured on multiple leaves per plant (i.e. per tiller). We reduced leaf-level observations to the plant-level, where if disease symptoms were observed on any leaf, we marked the disease as present for the plant. This yielded approximately 41,000 plant-level observations total.

Environmental Data

Weather station data was obtained from the North Durham Wastewater Treatment Facility (ID: DURH; North Carolina State Climate Office, 2025). This facility is approximately 10 km from Widener Field and was the closest station which recorded both atmospheric and soil conditions at all time points surveyed for disease. The final data set included 2,814 observations (daily mean value) of 32 climatic variables spanning April 24, 2017, to June 6, 2025. A full list of climatic variables, units, and descriptors is given in Table S1.

For analyses sensitive to multicollinearity, observations of the 32 climatic variables through time were reduced to four principal components (PCs) using the R function ‘prcomp’. Climatic variables were first centered and scaled to have a mean of zero and variance of one using the

151 ‘decostand’ function in R package vegan (method = “standardization”). The first four PCs
152 together explained 79% of the total variance in the climatic variables over time and explained
153 more variance than expected by chance under a broken stick model (Figure S1). For the week
154 preceding each fungal disease observation, we calculated the mean score for each principal
155 component. The average dynamics of these environmental PCs over the course of a year can be
156 found in Figure S2. Finally, we visualized the loadings of each climatic variable on the PCs to
157 infer relationships between the original 32 variables and 4 summary PCs (Figure S3).

158 *Random forest analysis to identify abiotic thresholds and predict annual pathogen dynamics*

159 Random forest models were developed using the R packages ‘ranger’ and ‘tidymodels’ (Kuhn &
160 Wickham, 2020; Wright & Ziegler, 2017). We trained binary classification models to predict the
161 presence or absence of each pathogen at the plant level. Predictor variables included the 7-day
162 average of climatic variables preceding each pathogen observation, the presence or absence of
163 other pathogens on the same plant, the mean prevalence of each pathogen in each plot at the
164 previous time point, month, and year. For example, the crown rust model had the following
165 structure:

166 $\text{Crown Rust}_{0|1,t} \sim \text{Anthracnose}_{0|1,t} + \text{Brown Patch}_{0|1,t} + \mu_{(t-7:t)}(\text{Climactic Variable}_i) +$
167 $\mu(\text{Crown Rust}_{t-1}) + \mu(\text{Anthracnose}_{t-1}) +$
168 $\mu(\text{Brown Patch}_{t-1}) + \text{Month} + \text{Year}$

169 To avoid temporal data leakage (where future information is inadvertently used to train or
170 evaluate a model and thus leads to unrealistically good performance), the data was split into
171 training and test sets based on year and month of observation. The first 80% of the data from

May 2017 – July 2022 were used as training sets, whereas the remaining 20% of the data from August 2022 – November 2024 was used as the test set.

Within the training data, we applied time series cross-validation using a sliding window approach. Each validation fold used the prior 12 months of data as training and tested on the subsequent month. This procedure generated a total of 36 sliding periods, though six of these contained validation sets with insufficient data (all spanning winter months, December – February) and were excluded, for a total of 30 final rolling windows. To address class imbalance, the data was up-sampled so that the number of pathogen presence cases equaled the number of absence cases in each temporal split.

For each random forest model, we grew 3,000 trees. The mtry hyperparameter, which controls the number of variables considered in each decision tree, was optimized via grid search over values from 4 to 32. The value minimizing the out-of-bag (OOB) error was selected for each model. The final 'mtry' parameter used for each model was 4 for anthracnose, 9 for crown rust, and 4 for brown patch.

Overall model performance was evaluated by comparing predictions to the held-out test set. We evaluated overall performance using accuracy, recall, precision, F-score, and kappa, which are described further in Table S2. Variable importance was assessed on the test data and is reported as permutation-based accuracy, indicating accuracy improvement for each predictor when compared to the permuted predictor. The relationship between climatic variables and focal pathogen presence were further explored using accumulated local effects (ALE) plots. We explored the ALE for the first three most important variables which were not collinear with any higher importance predictors (Pearson correlation coefficient < 0.8). ALE plots facilitate

194 understanding of how individual predictor variables affect the response of a machine learning
195 model, while controlling for cross-correlation and interactions among predictors. In our case, a
196 higher ALE value indicates that the model is more likely to predict pathogen presence for a given
197 parameter value than the model would predict on average.

198 *Survival analyses to estimate effects of earlier arriving pathogens on later arriving pathogens*

199 In 2018, individual plants were surveyed longitudinally (i.e., marked and tracked over time) for
200 disease. Using this data, a survival analysis was conducted to evaluate whether prior infection by
201 a different pathogen facilitated or antagonized subsequent infection, in the context of the
202 environment. For each pathogen, we constructed a Cox proportional hazards regression model,
203 with plant ID included as a clustering variable to account for repeated measures on the same
204 plant and environmental variables summarized by the first four principal components as
205 covariates. R package 'survival' was used to build models and 'survminer' was used to visualize
206 results (Kassambara et al., 2025; Therneau, 2026).

207 *Environmental niche overlap between pathogens*

208 We additionally calculated the degree of overlap between the pathogen's environmental niche
209 spaces using R package 'nicheROVER' (Lysy et al., 2023). Climatic variables were summarized
210 as PC scores for this analysis. Using this information, we calculated the pairwise probabilities
211 that a randomly drawn individual from pathogen A would be found within the interior 95% of
212 the niche region of pathogen B using function 'overlap' (alpha = .95).

213 *Bayesian hierarchical analysis to estimate the interaction between biotic and abiotic variables*

To complement the machine learning and survival models, we additionally constructed Bayesian regression models for our data. This Bayesian framework allowed us to account for hierarchical structure and quantify uncertainty in pathogen–pathogen and pathogen–environment interactions using the full plot-level dataset, which could not be easily addressed within the survival or random forest frameworks.

Separate models were built for each focal pathogen (Bernoulli family with logit link) using the R package 'brms' (Bürkner, 2021). Fixed effects included coinfecting pathogens, environmental variables summarized by the first four principal components, sampling month and year, and the focal pathogen's prevalence in the prior census at the same plot. Interaction terms between co-occurring pathogen presences and environmental variables were also included to assess environmentally dependent changes in the strength or direction of pathogen-pathogen interactions. A random effect was included to account for repeated resampling of plots over time. For example, the crown rust model had the following structure:

$$\begin{aligned}\text{Crown Rust}_{0|1,t} \sim\ & \mu(\text{Crown Rust}_{t-1}) + \text{Month} + \text{Year} + \text{EnvPC1} + \ldots + \text{EnvPC4} + \\ & \text{Anthracnose}_{0|1,t} + \text{Brown Patch}_{0|1,t} + \text{Anthracnose}_{0|1,t}:\text{Brown Patch}_{0|1,t} + \\ & \text{Anthracnose}_{0|1,t}:(\text{EnvPC1} + \ldots + \text{EnvPC4}) + \\ & \text{Brown Patch}_{0|1,t}:(\text{EnvPC1} + \ldots + \text{EnvPC4}) + \\ & (1 \mid \text{Plot ID})\end{aligned}$$

A weakly informative prior, *Normal*(0, 2), was applied to fixed effects to improve convergence and prevent overfitting of terms. Four chains were run per model with 6000 iterations total and 3000 burn in iterations. Posterior predictive checks and convergence diagnostics ($\hat{R} < 1.01$, ESS > 1000) confirmed adequate model across all models (Figures S4 – S6).

Results

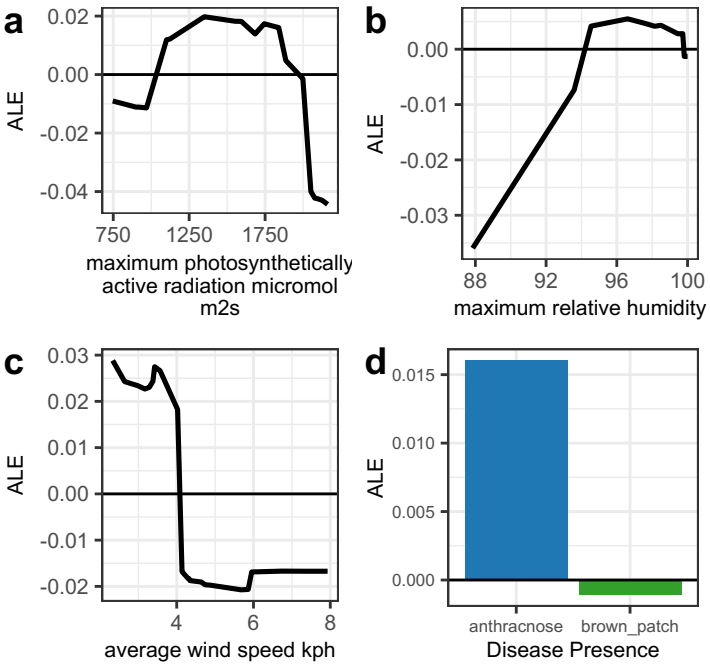

Figure 2. Random forest model results for crown rust (a-c) Three most important and non-collinear variables to predicting crown rust dynamics exhibit critical thresholds where the probability of disease observation rapidly changes, indicated by a change in ALE sign. (d) Anthracnose presence more strongly affected the probability of predicting crown rust for a given plant than brown patch presence.

Temporal random forest models well predict crown rust and brown patch dynamics

Crown rust, anthracnose, and brown patch dynamics were each best predicted by disparate environmental and biotic factors (Figures S7 – S9). Model performance metrics were comparable across brown patch (Accuracy = 0.844, Kappa = 0.197) and crown rust models (Accuracy = 0.808, Kappa = 0.285), indicating high predictive power. Conversely, the anthracnose model had lower predictive power (Accuracy = 0.456, Kappa = 0.0803) and was only nominally better than selecting classes based on random chance (Table S2). For each disease, we also constructed models with the 14-day and 30-day averages of environmental predictors leading up to the sampling date, rather than the 7-day average presented here. Models with these longer-term averages generally decreased model performance (Tables S3 – S4). For each model, we visualize the ALE of the three most important and non-collinear variables, but visualizations of the top 15 most important variables (after this threshold

there is a generally steep drop off in variable importance and most variables are colinear) are given in Figures S11 – S13.

Species-specific environmental thresholds and biotic interactions predict disease dynamics

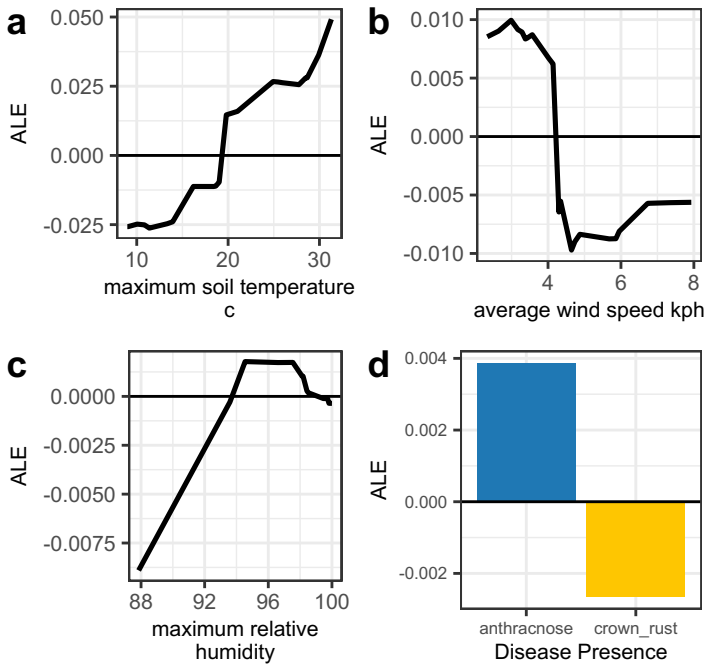

Figure 3. Random forest model results for brown patch. (a-c) Three most important and non-collinear variables to predicting brown patch dynamics exhibit critical thresholds where the probability of disease observation rapidly changes, indicated by a change in ALE sign. (d) Anthracnose co-occurrence more strongly affected the probability of predicting brown patch for a given plant than crown rust presence, although neither was in the top 15 most important variables to the model.

Crown rust dynamics were primarily driven by light-, humidity-, and wind-related variables, as well as anthracnose coinfection (Figure S7). However, the best predictor for crown rust presence on a host plant was its own lagged prevalence, indicating a high degree of temporal autocorrelation. Beyond this, crown rust was most likely to be observed when maximum photosynthetically active radiation ranged from 1000-2000 μmol/m^2/s, maximum daily relative humidity ranged from 93-99%, and average wind speeds were less than 4 kmph (Figure 2a-c). Crown rust presence at the plant-level was also positively associated with anthracnose coinfection (Figure 2d).

Brown patch dynamics were primarily predicted by temperature-related variables, followed by a smaller role for wind- and humidity-related predictors and lagged anthracnose and crown rust prevalence (Figure S8). Like crown rust, the best predictor of brown patch presence at the plant level was its own lagged prevalence in the host population, indicating a high degree of temporal autocorrelation. With respect to the environmental variables, brown patch was most likely to be observed when maximum soil temperature exceeded 18.8ºC, average wind speed was below 4 kmph, and maximum relative humidity ranged from 93.8%-98% (Figure 3a-c). While neither anthracnose or crown rust coinfection were in the 15 most important variables to the model (Figure 3d), the lagged prevalence of crown rust and anthracnose at the plot level generally increased the probability of observing brown patch (Figure S12).

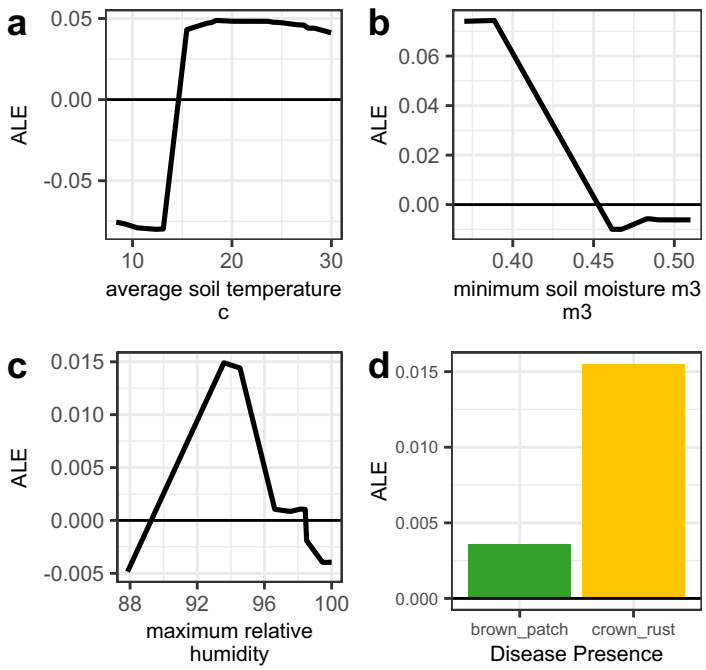

Figure 4. **Random forest model results for anthracnose**. (a-c) Three most important and non-collinear variables to predicting anthracnose dynamics exhibit critical thresholds where the probability of disease observation rapidly changes, indicated by a change in ALE sign. (d) Crown rust coinfection more strongly affected the probability of predicting crown rust for a given plant than brown patch presence.

While overall predictive power for anthracnose was low, its dynamics were primarily driven by temperature-, moisture-, and humidity-related variables (Figure S7). Anthracnose was most likely to be observed at the plant-level when daily average soil temperatures

exceeded 13ºC, daily average soil moisture was lower than 0.45 m^3/m^3, and daily maximum relative humidity ranged from 87.5-96.3% (Figure 4a-c). While neither crown rust or brown patch coinfection, or their previous prevalences in the host population, were among the most important variables in predicting anthracnose disease presence (Figure S7), crown rust coinfection increased the probability of observing anthracnose more than brown patch (Figure 4d).

Facilitation by early arriving pathogens can exceed the impact of abiotic factors on plant-level disease risk

Anthracnose, brown patch, and crown rust displayed substantial overlap in their climatic niches (Figure S10), making it possible that the apparent facilitative effects identified by our machine learning models result from shared environmental responses rather than biotic interactions. Moreover, these potential interactions may be masked in the full combined dataset from 2017–2024, since it contains repeated measurements at the plot level but lacks longitudinal tracking of individual plants in most years. To further investigate these factors, we utilized a subset of the full dataset from 2018 which contained longitudinal, individual-level data to test how prior pathogen infections and environmental variables influence the host individual's risk of acquiring each disease ("disease risk" for a focal disease) using Cox proportional hazards models.

For crown rust, climatic variables, prior infection by anthracnose and prior infection by brown patch all had significant effects on disease risk (Figure S14). Prior infection by anthracnose increased the risk of subsequent crown rust infection approximately 3-fold (Hazard ratio: 3.50, $p < 0.001$). Prior infection by brown patch nearly doubled the risk of subsequent crown rust infection (Hazard ratio: 1.61, $p < 0.001$). Additionally, we found that all four environmental PCs

significantly affected the risk of crown rust infection, albeit with smaller effect than prior infection by either anthracnose or brown patch.

For brown patch, while climatic variables and prior crown rust infection significantly impacted brown patch infection risk, prior anthracnose infection did not (Figure S15). We estimated that prior crown rust infection had the strongest effect on brown patch infection risk, roughly doubling it (Hazard ratio: 1.88, $p < 0.001$). Environmental PC1, PC3, and PC4 increased brown patch infection risk with smaller effect, likely representing the effects of temperature, wind, and soil moisture, respectively (Figure S3).

Finally, for anthracnose, climatic variables had the only significant effect on disease risk (Figure S16). In this model, the first four environmental PCs had a significant effect on the hazard ratio, with PC1 exhibiting the strongest effect (Hazard ratio = 1.44, $p < 0.001$). Temperature-related variables primarily contribute to PC1 (Figure S3), suggesting that higher temperatures were the strongest driver of anthracnose presence.

Environmental context can modulate the strength and direction of pathogen-pathogen interactions

Finally, we investigated potential interactions between the four environmental PCs and pathogen co-occurrence in driving focal disease presence. We built a Bayesian hierarchical model to predict the probability of observing the focal pathogen based on environmental PCs, time of year, pathogen coinfection, and the interaction of PCs and coinfection. By using these models, we were able to explicitly account for the repeated sampling of plots through time and retained sufficient power to look at interactions between environmental PCs and pathogen coinfection, lending us insights into how pathogen facilitation may strengthen or weaken under specific

environmental contexts. For each disease, we found significant interactions between environmental PCs and pathogen coinfection (Table S5). Here, we focus on these interaction terms, as they give insights not revealed by our other models, but full Bayesian model results can be found in Tables S6 – S8.

For crown rust, coinfection by either anthracnose or brown patch interacted with the effects of the abiotic predictors for all environmental PCs (Figure 5). The variable which was most predictive of crown rust presence was anthracnose coinfection ($\beta = 1.27, CI_{95\%} = [0.0832, 1.11]$), but this positive association of crown rust with anthracnose waned as PC1 values increased, representing warmer temperatures ($\beta = -0.0408, CI_{95\%} = [-0.0771, -0.00232]$). Increases in PC1 generally reduced the probability of crown rust infection ($\beta = -0.184, CI_{95\%} = [0.027, -0.235]$), but this association disappeared in plants coinfected by brown patch ($\beta = 0.181, CI_{95\%} = [0.131, 0.231]$). While PC2 did not have any base effects on the probability of crown rust infection, crown rust was positively associated with brown patch coinfection at low values of PC2; this association disappeared as PC2 values increased ($\beta = -0.182, CI_{95\%} = [-0.288, -0.0805]$), likely representing increasing humidity or decreasing winds. Increases along PC3, which was not well characterized by any individual environmental variable, decreased the probability of crown rust infection ($\beta = -0.116, CI_{95\%} = [-0.236, 0.000232]$). For plants with coinfection by either brown patch ($\beta = 0.119, CI_{95\%} = [-0.00178, 0.240]$) or anthracnose ($\beta = 0.0805, CI_{95\%} = [-0.0394, 0.207]$) this effect was generally reduced. Finally, increases in PC4 decreased the probability of crown rust infection ($\beta = -0.442, CI_{95\%} = [-0.531, -0.354]$). For plants coinfected by either brown patch ($\beta = 0.137, CI_{95\%} = [0.0419, 0.226]$) or anthracnose ($\beta = 0.237, CI_{95\%} = [0.155, 0.322]$), this negative relationship was more pronounced.

For brown patch, coinfection by either anthracnose or crown rust interacted with environmental PC1 and PC2 (Figure 6). Higher PC1 values, representing higher temperatures, increased the probability of brown patch infection ($\beta = 0.258, CI_{95\%} = [0.188, 0.327]$), and this effect was intensified by crown rust coinfection ($\beta = 0.0629, CI_{95\%} = [0.0151, 0.112]$). Conversely, coinfection by anthracnose decreased the probability of brown patch infection at high PC1 values ($\beta = -0.215, CI_{95\%} = [-0.259, -0.169]$). Higher PC2 values, representing increased humidity or lower winds, also increased the probability of brown patch infection ($\beta = 0.362, CI_{95\%} = [0.266, 0.456]$). Coinfection by either anthracnose ($\beta = -0.284, CI_{95\%} = [-0.382, -0.187]$) or crown rust ($\beta = -0.163, CI_{95\%} = [-0.256, -0.0700]$) lessened this effect, decreasing the probability of brown patch infection at higher PC2 values. While higher PC3 ($\beta = 0.125, CI_{95\%} = [-0.00486, 0.256]$) and lower PC4 ($\beta = -0.203, CI_{95\%} = [-0.303, -0.107]$) values increased the probability of brown patch infection, they did not interact with anthracnose and crown rust coinfection.

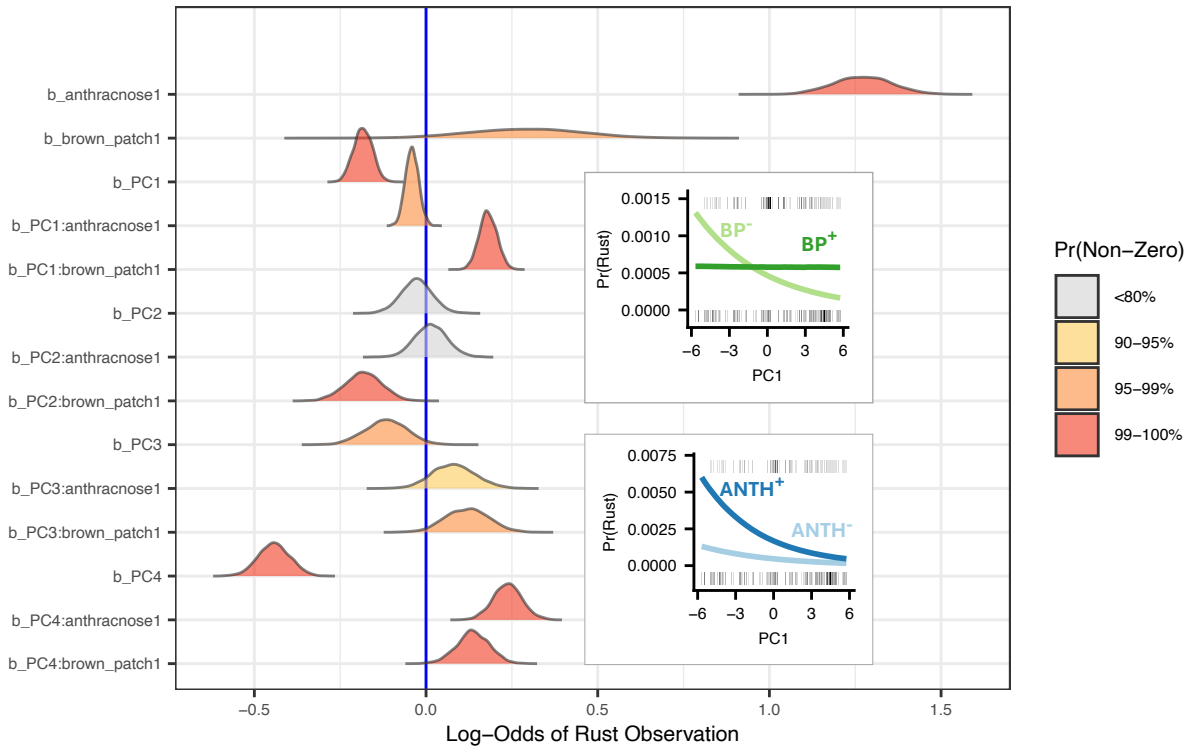

Figure 5. Bayesian posterior estimates for fixed effects of environmental PCs, pathogen coinfection, and their interactions in the crown rust model. Color indicates the posterior probability that an effect is nonzero. Insets highlight conditional effects of selected PC × coinfection interactions, with ANTH+/ANTH− denoting presence/absence of anthracnose and BP+/BP− denoting presence/absence of brown patch. Rugs along the top of the inset give the density of positive crown rust observations (CR+) with respect to the environmental PC, while rugs along the bottom of the inset give the density of crown rust absence (CR-) observations. Conditional effects of all pathogen x PC interactions can be found in Figure S17.

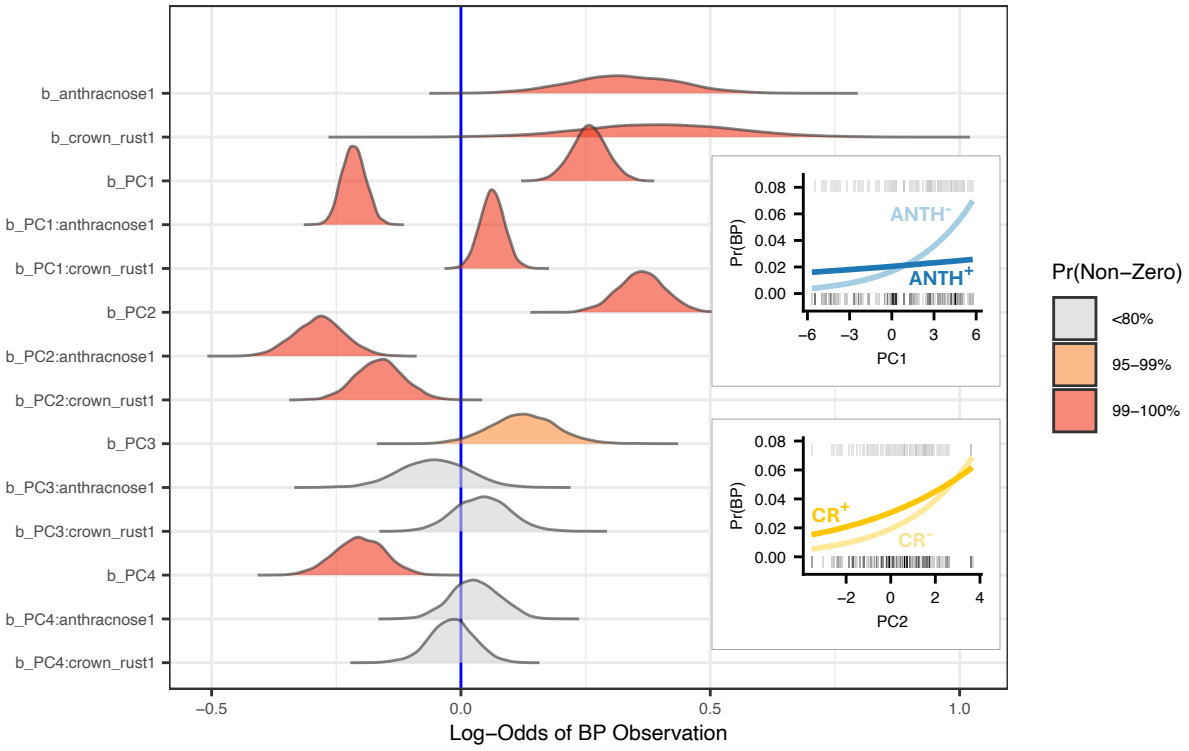

Figure 6. Bayesian posterior estimates for fixed effects of environmental PCs, pathogen coinfection, and their interactions in the brown patch model. Color indicates the posterior probability that an effect is nonzero. Insets highlight conditional effects of selected PC × coinfection interactions, with ANTH+/ANTH− denoting presence/absence of anthracnose and

CR+/CR− denoting presence/absence of crown rust. Rugs along the top of the inset give the density of positive brown patch (BP+) observations with respect to the environmental PC, while rugs along the bottom of the inset give the density of brown patch absence (BP-) observations. Conditional effects of all pathogen x PC interactions can be found in Figure S18.

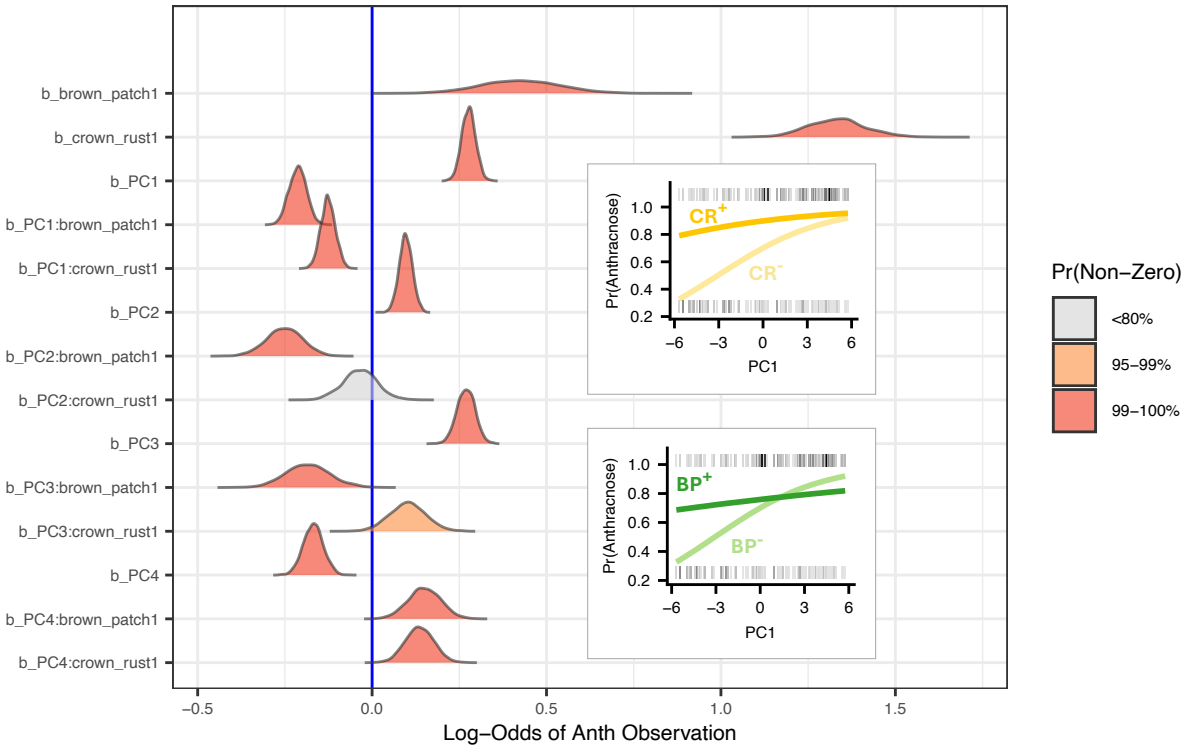

Figure 7. Bayesian posterior estimates for fixed effects of environmental PCs, pathogen coinfection, and their interactions in the anthracnose model. Color indicates the posterior probability that an effect is nonzero. Insets highlight conditional effects of selected PC × coinfection interactions, with BP+/BP− denoting presence/absence of brown patch and CR+/CR− denoting presence/absence of crown rust. Rugs along the top of the inset give the density of positive anthracnose observations (ANTH+) with respect to the environmental PC, while rugs along the bottom of the inset give the density of anthracnose absence observations (ANTH-). Conditional effects of all pathogen x PC interactions can be found in Figure S19.

Finally, for anthracnose, coinfection by crown rust or brown patch generally lessened the association between abiotic predictors and anthracnose infection (Figure 7). Higher environmental PC1 values, indicating higher temperatures (Figure S3) were positively associated with anthracnose infection ($\beta = 0.277, CI_{95\%} = [0.238, 0.315]$). Plants without crown rust or brown patch coinfections were more tightly associated with changes along PC1. PC2 was positively associated with anthracnose infection ($\beta = 0.0966, CI_{95\%} = [0.0591, 0.134]$), where high values represent high humidity variables. Brown patch coinfection reduced the effects of PC2 on anthracnose infection ($\beta = -0.251\ CI_{95\%} = [-0.354, -0.150]$). Increased PC3 values, which likely represents a combination of wind, moisture, and sunlight variables, also increased the probability of anthracnose infection ($\beta = 0.269, CI_{95\%} = [0.219, 0.319]$). For plants coinfected by brown patch, this effect was lessened ($\beta = -0.184, CI_{95\%} = [-0.303, -0.05]$), while coinfection by crown rust intensified this effect ($\beta = 0.102, CI_{95\%} = [-0.0004, 0.204]$). Finally, increases in PC4, likely representing decreased soil moisture, decreased the probability of anthracnose infection ($\beta = -0.169, CI_{95\%} = [-0.224, -0.110]$). Brown patch ($\beta = 0.150, CI_{95\%} = [0.0593, 0.240]$) or crown rust ($\beta = 0.136, CI_{95\%} = [0.0586, 0.215]$) coinfection both lessened the negative effects of PC4 on anthracnose infection probability.

Discussion

From this study, we draw two major conclusions, one ecological and one methodological. Ecologically, we found that that pathogen coinfection can influence disease trajectories at a similar strength as abiotic factors. We further showed that the magnitude and direction of these interactions can vary with environmental context. Methodologically, through using temporally explicit random forest models, we were able to generate accurate predictions of field epidemics and identify critical environmental thresholds for two of three focal diseases. We then combined

429 these machine learning models with Bayesian hierarchical models and survival analyses to gain a
430 better mechanistic understanding of epidemic dynamics. Taken together, we provide an
431 integrated framework for understanding and forecasting the dynamics of multi-pathogen
432 systems.

433 Accurate prediction of field disease dynamics may require explicitly accounting for coinfections.
434 When multiple pathogens circulate simultaneously, single-pathogen frameworks can be
435 qualitatively misleading because they fail to capture effects of pathogen–pathogen interactions
436 (Strauss et al., 2021; Dutt et al., 2022; Garin et al., 2018; Susi et al., 2015). Our results suggest
437 that some pathogens can increase host susceptibility to infection by other, later-arriving
438 pathogens and thereby alter epidemic trajectories, consistent with other evidence that historical
439 contingency can shape disease outcomes (Halliday et al., 2020). When early arriving pathogens
440 modify host susceptibility or local pathogen communities, environmental conditions favoring
441 one pathogen may propagate through the system to influence later epidemics of other pathogens
442 in unexpected ways.

443 Further, pathogen-pathogen interactions may vary based on environmental context. Such context
444 dependence was common in our data (Figure 5 - Figure 7, Figure S17 - Figure S19). However,
445 these apparent context-dependent interactions may alternatively reflect phenological patterns,
446 rather than true pathogen–pathogen interactions. This challenge of distinguishing true
447 interactions from temporal co-occurrence patterns motivated our inclusion of month as a
448 covariate in the Bayesian and random forest models, allowing us to assess the predictive power
449 added by including environmental PCs as well as time of season per se (i.e. month). In some
450 models, month was a stronger predictor than environmental PCs, reaffirming that seasonal

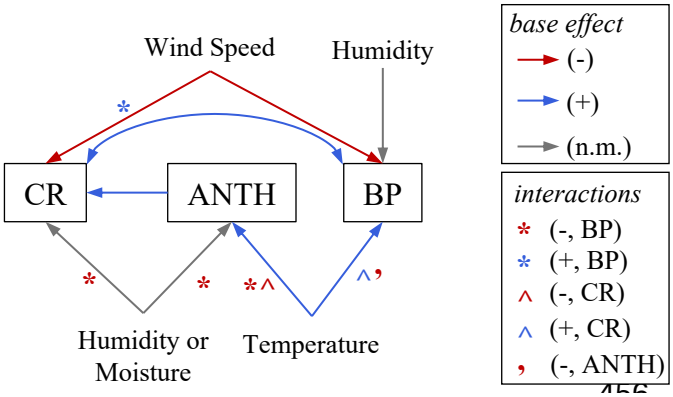

Figure 8. Inferred environmental drivers and biotic interactions for crown rust (CR), anthracnose (ANTH), and brown patch (BP), based on consensus across the three models. Humidity and temperature are well-established fungal disease drivers; wind speed likely reflects dispersal (CR) or a correlated unmeasured variable such as thunderstorm intensity (CR, BP). For each variable, we report the consensus direction of effect (positive, negative, or non-monotonic) on each pathogen. Humidity and moisture are classified as non-monotonic (n.m.), indicating optimal ranges for symptom expression, while other variables show threshold or linear responses. Asterisks/annotations indicate inferred directions of coinfection-environment interactions.

timing, such as the arrival of crown rust from alternate hosts or shifts in plant community composition, plays a substantial role in disease emergence and persistence.

Despite their apparent importance, incorporating multi-pathogen dynamics into predictive models can present substantial challenges. Coinfection processes introduce nonlinearity and context dependence which can be poorly captured by traditional approaches. We found random forest models to be a robust and flexible framework to address these challenges, building upon existing work leveraging them in single pathogen plant systems (Alves et al., 2025; Shah et al., 2023) to generate predictions (>80% accuracy) for two of our three focal pathogens. Further, we recovered biologically relevant environmental thresholds from noisy environmental data for all three of our focal pathogens.

The identified environmental thresholds align with variables known to be important to fungal phenology in general (e.g., temperature, moisture; Delmas et al., 2024) and with the life cycles of our focal fungal diseases. For example, *Colletotrichum cereale*, a causal agent of anthracnose, favors high humidity conditions with optimal temperatures of 21-27°C (*Turfgrass Diseases*, n.d.), coarsely aligning with the results from our random forest model. Paralleling this, the recovery of maximum photosynthetically active radiation as an important variable to crown rust presence likely reflects the host plant's growing season and *P. coronata's* obligate biotrophy (Butler & Kerns, 2019b), whereas the increased probability of observing *P. coronata* at lower, but non-zero, wind speeds may reflect ideal conditions for dispersal. Finally, *Rhizoctonia solani*, the causal agent of brown patch, overwinters in the soil as a saprobe and we recovered soil temperature to be the most important environmental driver of brown patch presence in the field (Butler & Kerns, 2019a).

As well as mechanistic insight, this integrated approach could support ecological forecasting by combining near-term predictions from machine learning models with scenario-based forecasts from mechanistic models (Ye et al., 2025). This would provide both immediate management guidance and longer-term planning horizons under changing environmental conditions.

Random forests excel at prediction but often do not fully leverage experimental design details aimed to give deeper mechanistic insight. We therefore complemented this approach with survival models that explicitly leveraged longitudinal plant-level data, allowing us to further distinguish facilitation from shared environmental niche space. From the combined modeling results, we generated a hypothesis integrating the most important environmental drivers identified from our random forest models, the pathogen-pathogen interactions from our survival analyses, and the environment-by-pathogen interactions from our Bayesian models (Figure 8).

Together, this hypothesized framework identifies interaction pathways most likely to mediate disease dynamics across environmental gradients.

Conclusions and Future Directions

Advancing ecological forecasting in multi-pathogen systems will require the integration of complementary modeling frameworks. Machine learning models can identify dominant predictors and nonlinear responses, while mechanistic models provide better insight into interaction strength and direction. Ensemble approaches, which combine predictions from multiple models, may offer a similarly powerful strategy for multi-pathogen systems (Chauvin et al., 2025; Mao et al., 2025). In our study, longitudinal data was key to inferences about pathogen species interactions, but its limitation to a short time period limited our ability to incorporate survival analyses into a formal ensemble framework. Future studies pairing finer-scale environmental measurements (e.g., at the plot level) with long-term, plant-level longitudinal data could enable more fully integrated forecasts that balance predictive accuracy with mechanistic understanding of how coinfection dynamics shift across environmental gradients. Developing such forecasting capacity will be critical for proactive disease management as climate change alters both pathogen phenology and the environmental context in which pathogen interactions occur.

Acknowledgements

This project would not have been possible without the many technicians who contributed to fieldwork from 2017-2024. This study was supported by NSF-USDA grant 2016-67013-25762 and NSF grant DEB-2308472 to CEM.

References

Altizer, S., Dobson, A., Hosseini, P., Hudson, P., Pascual, M., & Rohani, P. (2006). Seasonality and the dynamics of infectious diseases. *Ecology Letters*, *9*(4), 467–484. https://doi.org/10.1111/j.1461-0248.2005.00879.x

Alves, K. S., Shah, D. A., Dillard, H. R., Ponte, E. M. D., & Pethybridge, S. J. (2025). Safer and Smarter: Leveraging Interpretation-Guided Modeling and Data Merging of Disease and Environmental Data for Plant Disease Risk Prediction. *Phytopathology®*, *115*(10), 1329–1343. https://doi.org/10.1094/PHYTO-01-25-0008-FI

Bürkner, P.-C. (2021). Bayesian Item Response Modeling in R with brms and Stan. *Journal of Statistical Software*, *100*(5), 1–54. https://doi.org/10.18637/jss.v100.i05

Butler, L., & Kerns, J. (2019a). *Brown Patch in Turf*. NC State Extension Publications. https://content.ces.ncsu.edu/brown-patch-in-turf

Butler, L., & Kerns, J. (2019b, December 16). *Rust in Turf*. NC State Extension Publications. https://content.ces.ncsu.edu/rust-in-turf

Butler, L., & Kerns, J. (2019c, December 19). *Anthracnose in Turf*. NC State Extension Publications. https://content.ces.ncsu.edu/anthracnose-in-turf

Chauvin, D., Gabriel, E., Martinetti, D., Papaïx, J., Martinez, C., Geniaux, G., Joudelat, F., & Soubeyrand, S. (2025). Factor Analysis and Prediction of Disease Risk Based on Large Ensembles of Models: Application to Virus Yellows in Sugar Beet. *Phytopathology®*, *115*(10), 1368–1381. https://doi.org/10.1094/PHYTO-01-25-0014-FI

Clay, P. A., Duffy, M. A., & Rudolf, V. H. W. (2020). Within-host priority effects and epidemic timing determine outbreak severity in co-infected populations. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1922), 20200046. https://doi.org/10.1098/rspb.2020.0046

Crouch, J. A., & Beirn, L. A. (2009). Anthracnose of cereals and grasses. *Fungal Diversity*, *39*, 19–44.

Delmas, C. E. L., Bancal, M.-O., Leyronas, C., Robin, M.-H., Vidal, T., & Launay, M. (2024). Monitoring the phenology of plant pathogenic fungi: Why and how? *Biological Reviews*, *99*(3), 1075–1084. https://doi.org/10.1111/brv.13058

Dutt, A., Andrivon, D., & Le May, C. (2022). Multi-infections, competitive interactions, and pathogen coexistence. *Plant Pathology*, *71*(1), 5–22. https://doi.org/10.1111/ppa.13469

Farnsworth, M. L., Hoeting, J. A., Hobbs, N. T., & Miller, M. W. (2006). Linking Chronic Wasting Disease To Mule Deer Movement Scales: A Hierarchical Bayesian Approach. *Ecological Applications*, *16*(3), 1026–1036. https://doi.org/10.1890/1051-0761(2006)016%255B1026:LCWDTM%255D2.0.CO;2

Fukami, T. (2015). Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. *Annual Review of Ecology, Evolution, and Systematics*, *46*(1), 1–23. https://doi.org/10.1146/annurev-ecolsys-110411-160340

Garin, G., Pradal, C., Fournier, C., Claessen, D., Houlès, V., & Robert, C. (2018). Modelling interaction dynamics between two foliar pathogens in wheat: A multi-scale approach. *Annals of Botany*, *121*(5), 927–940. https://doi.org/10.1093/aob/mcx186

Gibson, G. J., Kleczkowski, A., & Gilligan, C. A. (2004). Bayesian analysis of botanical epidemics using stochastic compartmental models. *Proceedings of the National Academy of Sciences*, *101*(33), 12120–12124. https://doi.org/10.1073/pnas.0400829101

Grunberg, R. L., Halliday, F. W., O'Keeffe, K. R., Joyner, B. N., Heckman, R. W., & Mitchell, C. E. (2025). Disease epidemics and species interactions: A manipulation of seasonal establishment of fungal diseases in an old field. *Ecology. 106: E70051.*, *106*, e70051.

Halliday, F. W., Penczykowski, R. M., Barrès, B., Eck, J. L., Numminen, E., & Laine, A.-L. (2020). Facilitative priority effects drive parasite assembly under coinfection. *Nature Ecology & Evolution*, *4*(11), 1510–1521. https://doi.org/10.1038/s41559-020-01289-9

Halliday, F. W., Umbanhowar, J., & Mitchell, C. E. (2017). Interactions among symbionts operate across scales to influence parasite epidemics. *Ecology Letters*, *20*(10), 1285–1294. https://doi.org/10.1111/ele.12825

Hassell, J. M., Newbold, T., Dobson, A. P., Linton, Y.-M., Franklinos, L. H. V., Zimmerman, D., & Pagenkopp Lohan, K. M. (2021). Towards an ecosystem model of infectious disease. *Nature Ecology & Evolution*, *5*(7), 907–918. https://doi.org/10.1038/s41559-021-01454-8

Kassambara, A., Kosinski, M., Biecek, P., & Fabian, S. (2025). *survminer: Drawing Survival Curves using "ggplot2"* (Version 0.5.1) [Computer software]. https://cran.r-project.org/web/packages/survminer/index.html

Kick, D. R., & Washburn, J. D. (2023). Ensemble of best linear unbiased predictor, machine learning and deep learning models predict maize yield better than each model alone. *In Silico Plants*, *5*(2), diad015. https://doi.org/10.1093/insilicoplants/diad015

Kuhn, M., & Wickham, H. (2020). *Tidymodels: A collection of packages for modeling and machine learning using tidyverse principles.* https://www.tidymodels.org

Lysy, M., Stasko, A. D., & Swanson, H. K. (2023). *nicheROVER: Niche Region and Niche Overlap Metrics for Multidimensional Ecological Niches* (Version 1.1.2) [Computer software]. https://cran.r-project.org/web/packages/nicheROVER/index.html

Mao, R., Shi, Y., Zhu, T., Zhou, J., Wang, J., Zhao, L., Liu, M., Xu, X., & Hu, X. (2025). Machine Learning Approaches for Predicting Fusarium Head Blight Epidemic Levels

from Climatological Time Series Features. *Phytopathology®*, *115*(10), 1344–1356. https://doi.org/10.1094/PHYTO-06-25-0221-FI

Molnar, C. (2025). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable* (3rd ed.). https://christophm.github.io/interpretable-ml-book

Natural Resources Conservation Service, U. S. D. of A. (2001, March 22). *Tall Fescue*. PLANTS Database. https://plants.usda.gov

Ogoshi, A. (1987). Ecology and Pathogenicity of Anastomosis and Intraspecific Groups of Rhizoctonia Solani Kuhn. *Annual Review of Phytopathology*, *25*(Volume 25,), 125–143. https://doi.org/10.1146/annurev.py.25.090187.001013

O'Keeffe, K. R., Simha, A., & Mitchell, C. E. (2021). Indirect interactions among co-infecting parasites and a microbial mutualist impact disease progression. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1956), 20211313. https://doi.org/10.1098/rspb.2021.1313

Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction. *Ecosystems*, *9*(2), 181–199. https://doi.org/10.1007/s10021-005-0054-1

Scherm, H., & Ojiambo, P. S. (2004). Applications of Survival Analysis in Botanical Epidemiology. *Phytopathology®*, *94*(9), 1022–1026. https://doi.org/10.1094/PHYTO.2004.94.9.1022

Shah, D. A., De Wolf, E. D., Paul, P. A., & Madden, L. V. (2023). Into the Trees: Random Forests for Predicting Fusarium Head Blight Epidemics of Wheat in the United States. *Phytopathology®*, *113*(8), 1483–1493. https://doi.org/10.1094/PHYTO-10-22-0380-R

Strauss, A. T., Bowerman, L., Porath-Krause, A., Seabloom, E. W., & Borer, E. T. (2021). Mixed infection, risk projection, and misdirection: Interactions among pathogens alter links between host resources and disease. *Ecology and Evolution*, *11*(14), 9599–9609. doi.org/10.1002/ece3.7781

Suh, D. C., Lance, S. L., & Park, A. W. (2024). Abiotic and biotic factors jointly influence the contact and environmental transmission of a generalist pathogen. *Ecology and Evolution*, *14*(8), e70167. doi.org/10.1002/ece3.70167

Susi, H., Barrès, B., Vale, P. F., & Laine, A.-L. (2015). Co-infection alters population dynamics of infectious disease. *Nature Communications*, *6*(1), 5975. doi.org/10.1038/ncomms6975

Therneau, T. M. (2026). *A Package for Survival Analysis in R*. https://CRAN.R-project.org/package=survival

Turfgrass Diseases: Anthracnose (Causal fungus: Colletotrichum cereale). (n.d.). Retrieved December 11, 2025, from https://extension.psu.edu/turfgrass-diseases-anthracnose-causal-fungus-colletotrichum-cereale

Vannette, R. L., & Fukami, T. (2014). Historical contingency in species interactions: Towards niche-based predictions. *Ecology Letters*, *17*(1), 115–124. doi.org/10.1111/ele.12204

Wright, M. N., & Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*, *77*(1), 1–17. https://doi.org/10.18637/jss.v077.i01

Ye, Y., Pandey, A., Bawden, C. Sumsuzzman, D.M., Rajput, R., Shoukat, A., Singer, B., Moghadas, S. M., Galvani, A. P. (2025) Integrating artificial intelligence with mechanistic epidemiological modleing: a scoping review of opportunities and challenges. *Nature Communications*, *16*(581). doi: 10.1038/s41467-024-55461-x

Supplemental Figures

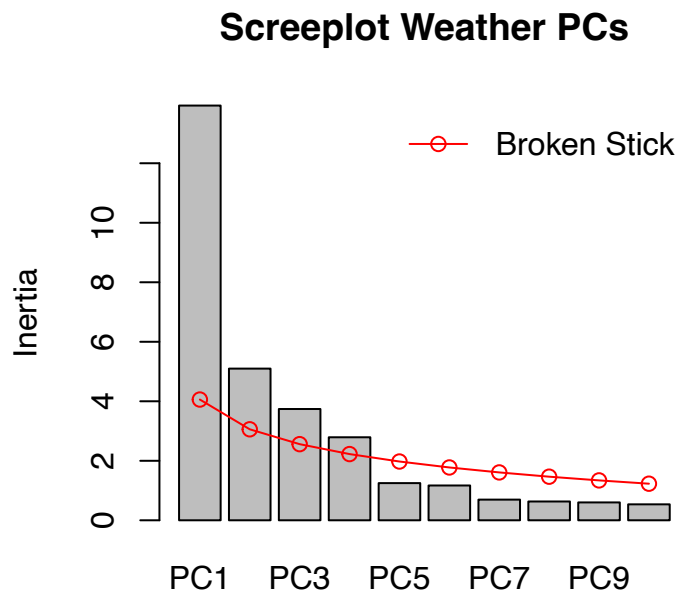

Figure S1. Scree plot of environmental PCs. Red line gives the broken stick null model (amount of variation expected to be explained by chance for each PC). The first four PCs explain more variation than expected by chance and were retained for subsequent analyses.

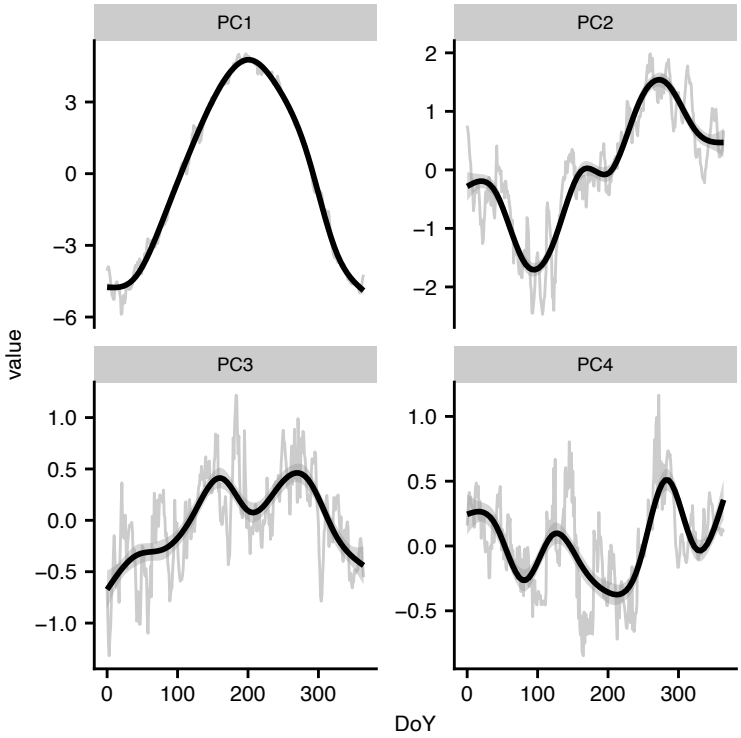

Figure S2. Figure SX. Average dynamics of each environmental PC over the course of a year. DoY represents "Day of the Year" on the x-axis, with day 0 = January 1. Trendline is a general additive model fit to the data with a cubic regression spline and formula y ~ s(x, bs = "cs").

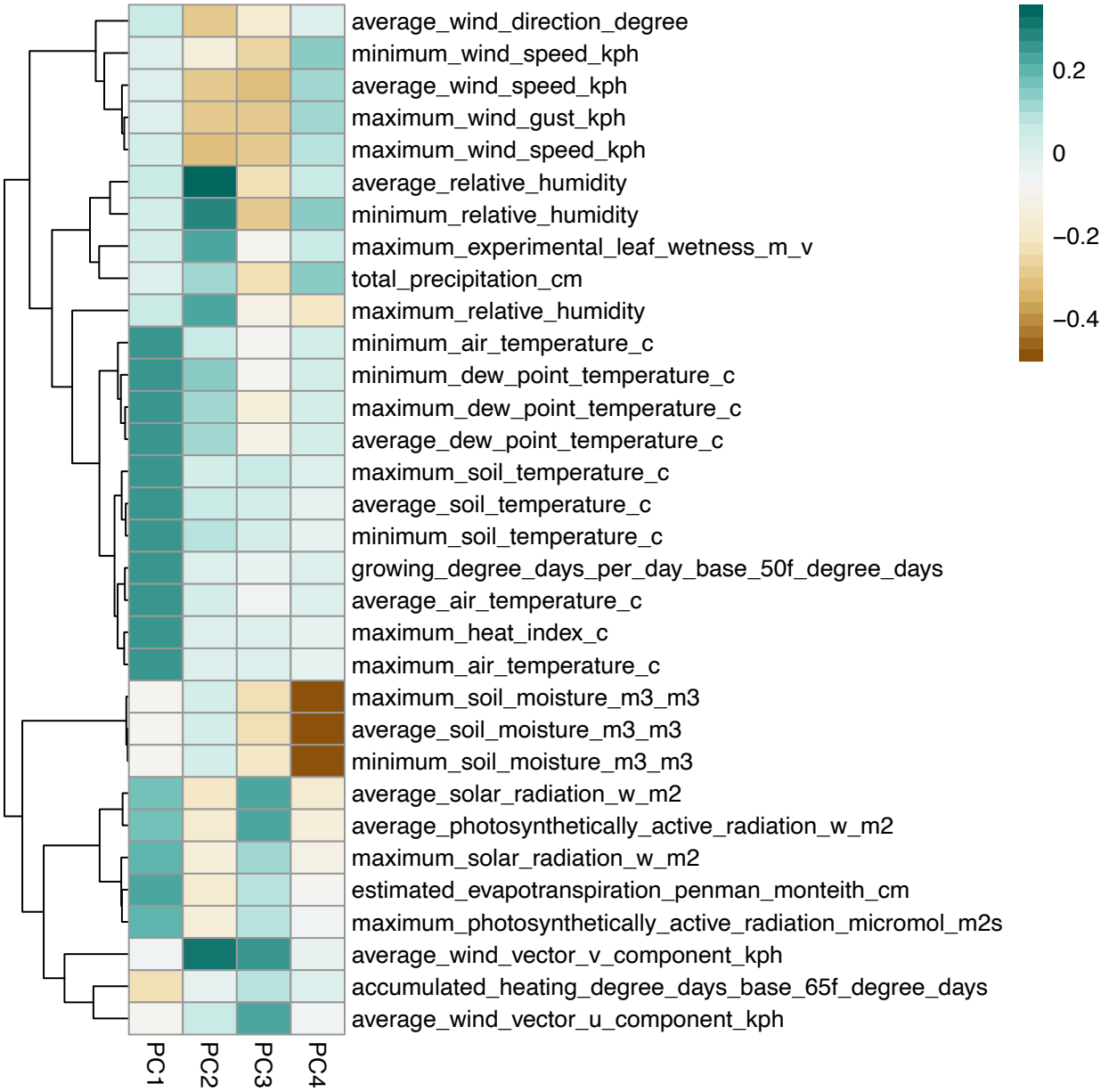

Figure S3. Heatmap of climatic variable loadings on the first four principal components, representing each variable's contribution to the components. Rows are clustered to highlight variables with similar loading patterns across components. Color intensity and hue indicate the magnitude and direction of each loading.

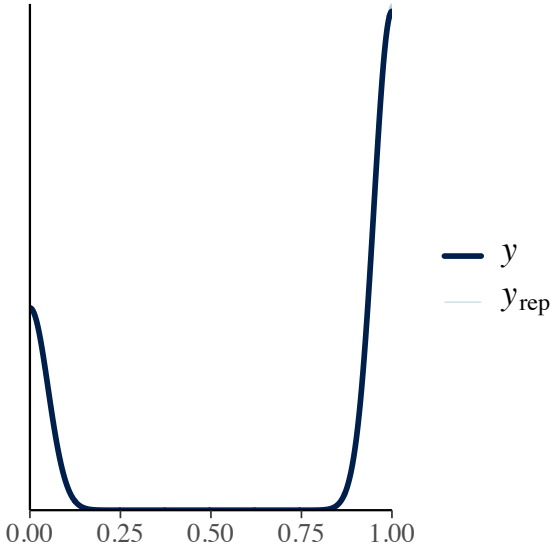

Figure S4. Posterior predictive check for anthracnose 'brms' model using a density overlay from 50 posterior draws. The thick line shows the density of the observed data, while the thin, semi-transparent lines represent densities of replicated datasets simulated from the posterior predictive distribution. Good agreement between observed and simulated densities indicates that the model adequately captures the overall distribution of the data; here, thin blue lines are difficult to see as they almost completely overlap with expected values.

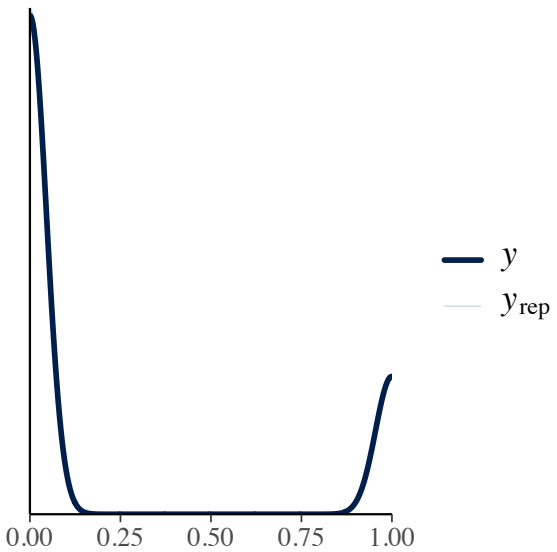

Figure S5 Posterior predictive check for crown rust 'brms' model using a density overlay from 50 posterior draws. The thick line shows the density of the observed data, while the thin, semi-transparent lines represent densities of replicated datasets simulated from the posterior predictive distribution. Good agreement between observed and simulated densities indicates that the model adequately captures the overall distribution of the data; here, thin blue lines are difficult to see as they almost completely overlap with expected values.

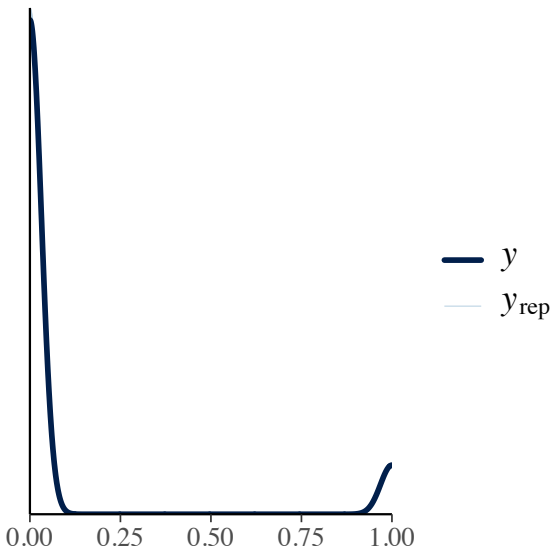

Figure S6. Posterior predictive check for brown patch 'brms' model using a density overlay from 50 posterior draws. The thick line shows the density of the observed data, while the thin, semi-transparent lines represent densities of replicated datasets simulated from the posterior predictive distribution. Good agreement between observed and simulated densities indicates that the model adequately captures the overall distribution of the data; here, thin blue lines are difficult to see as they almost completely overlap with expected values.

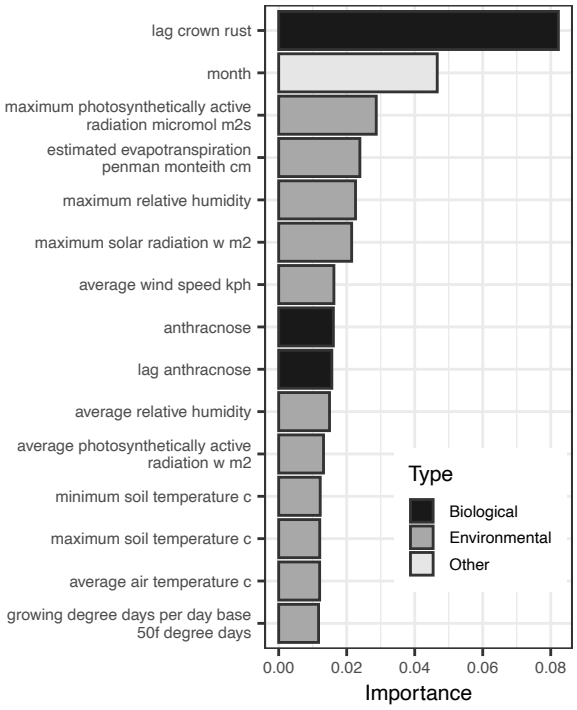

Figure S7. Top fifteen most important variables in predicting crown rust dynamics from random forest model.

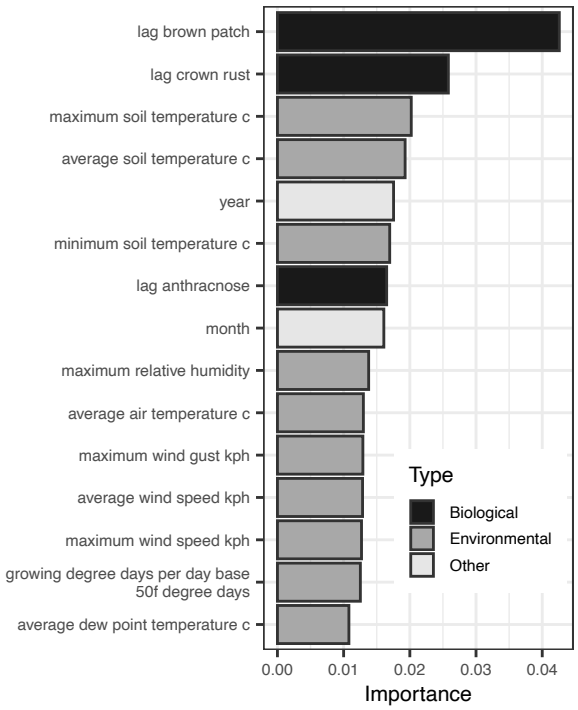

Figure S8. Top fifteen most important variables in predicting brown patch disease state from random forest model.

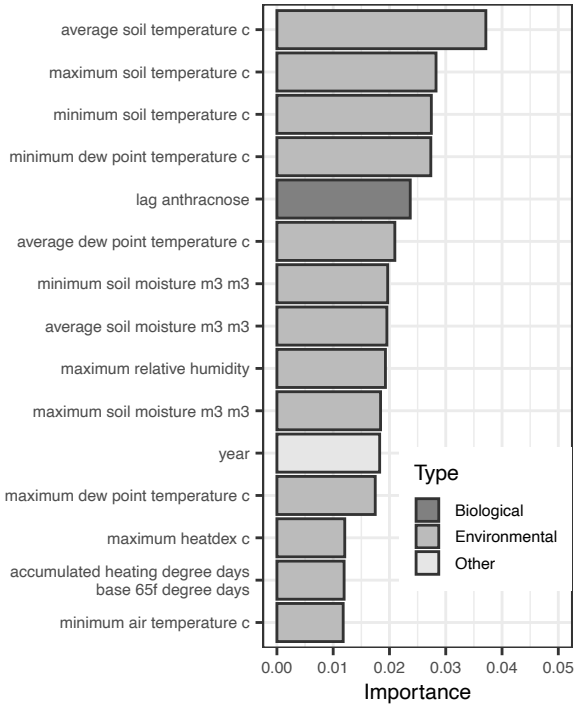

Figure S9. Top fifteen most important variables to predict anthracnose disease state from random forest model.

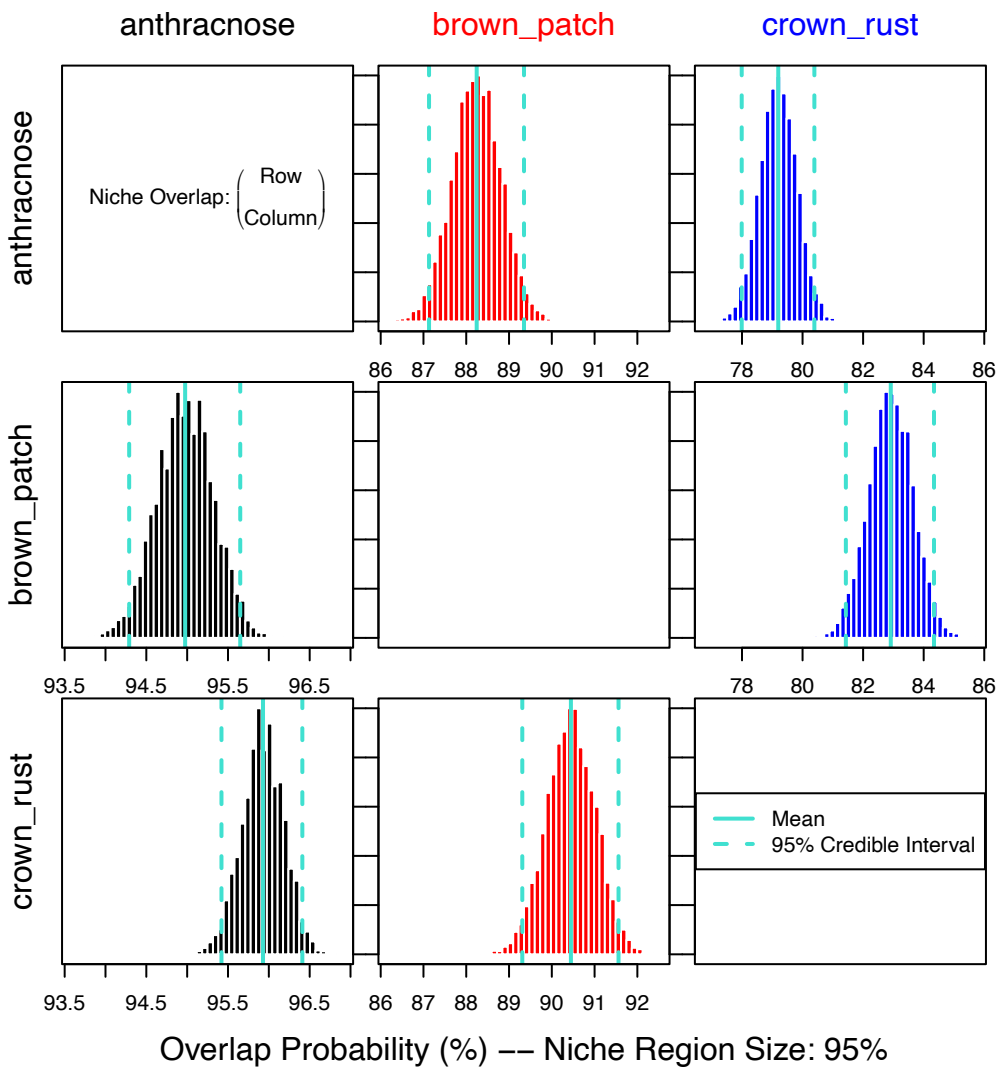

Figure S10. Niche overlap among anthracnose, brown patch, and crown rust, calculated from environmental principal components. Rows represent the focal pathogen, and columns indicate the probability that the focal pathogen occurs within the core 95% of another pathogen's environmental niche. For example, there is an 88.2% probability that anthracnose occurs within 95% of the environmental niche of brown patch.

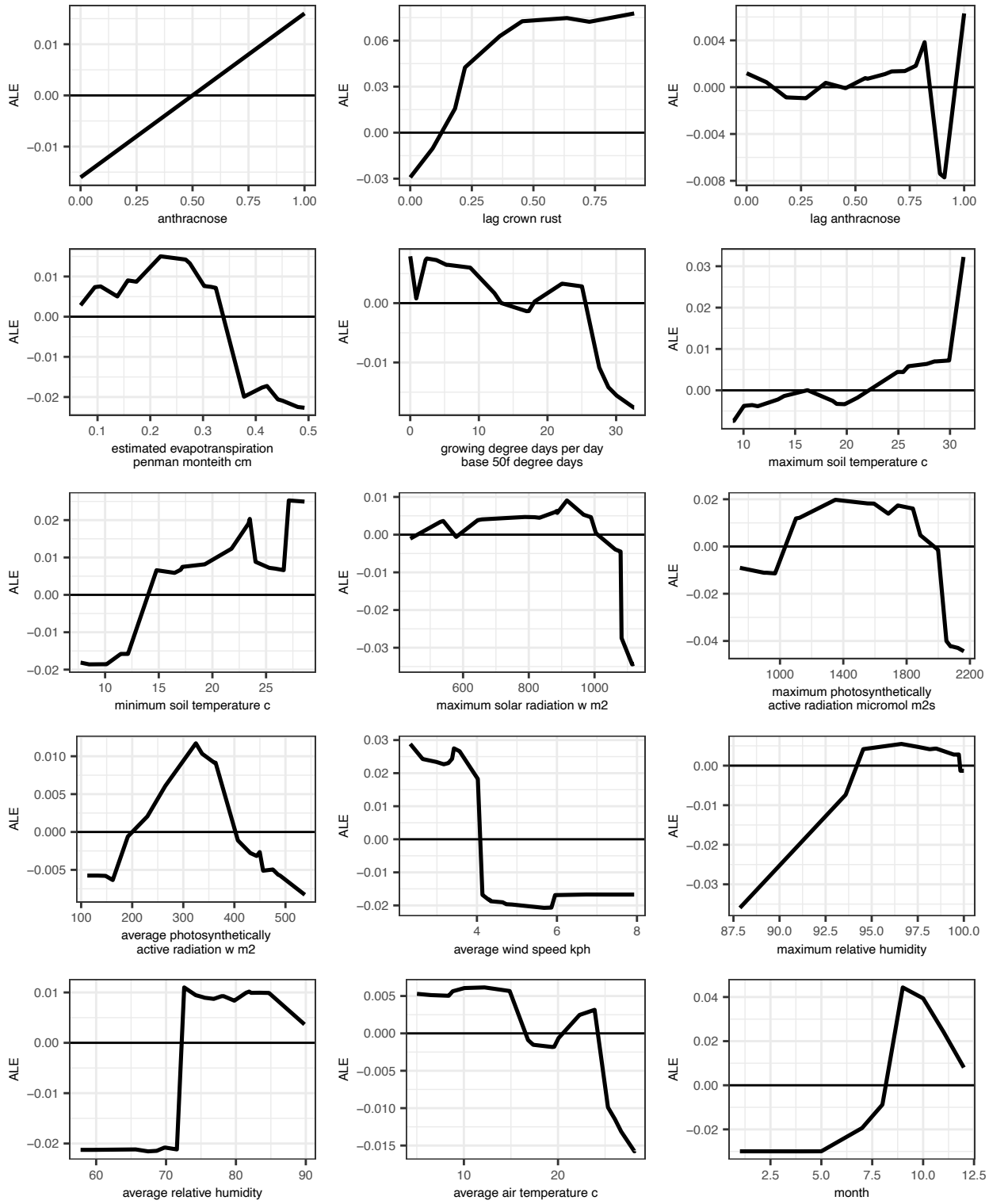

Figure S11. ALE plots for top 15 most important variables in crown rust model.

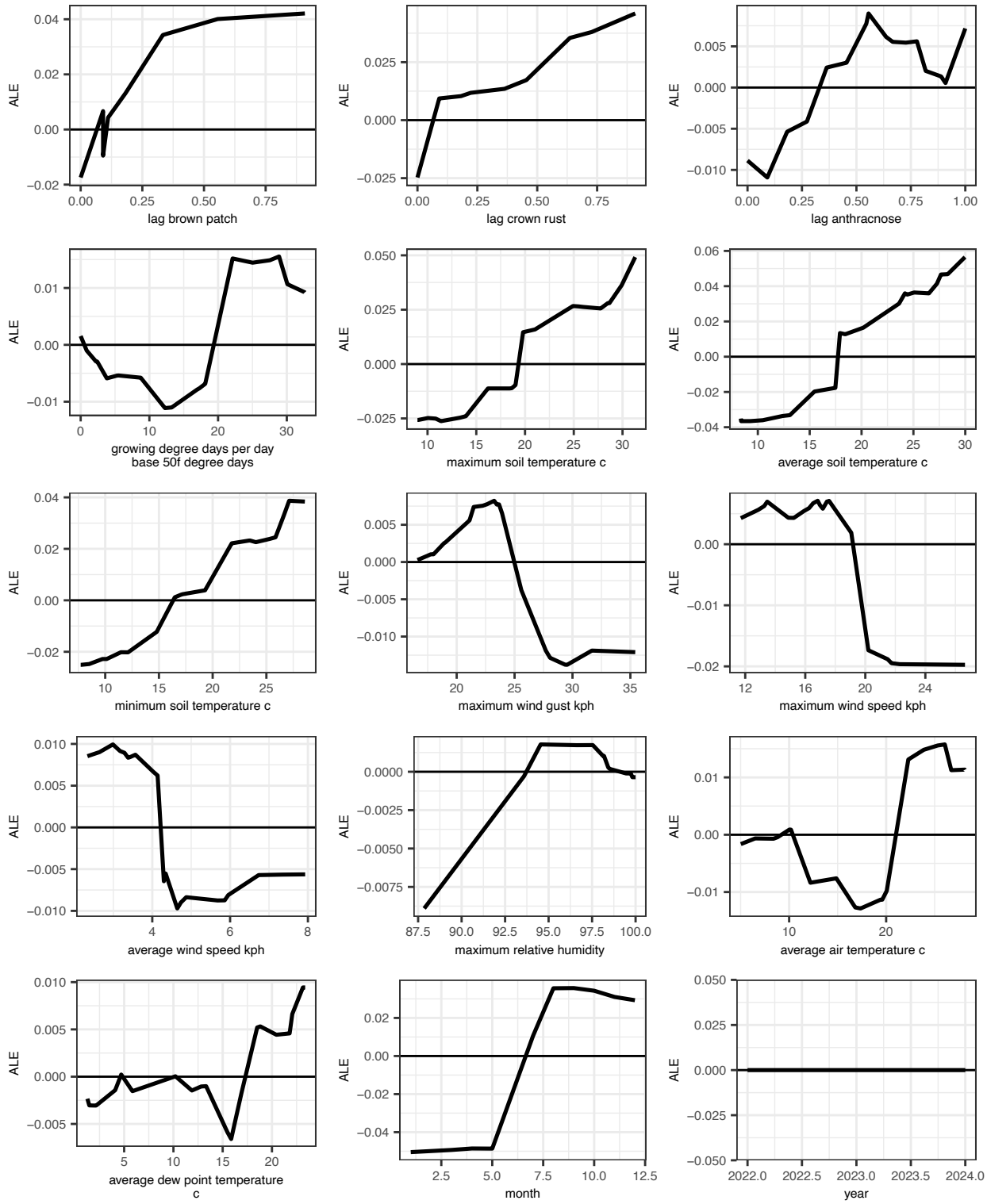

Figure S12. ALE plots for top 15 most important variables in brown patch random forest model.

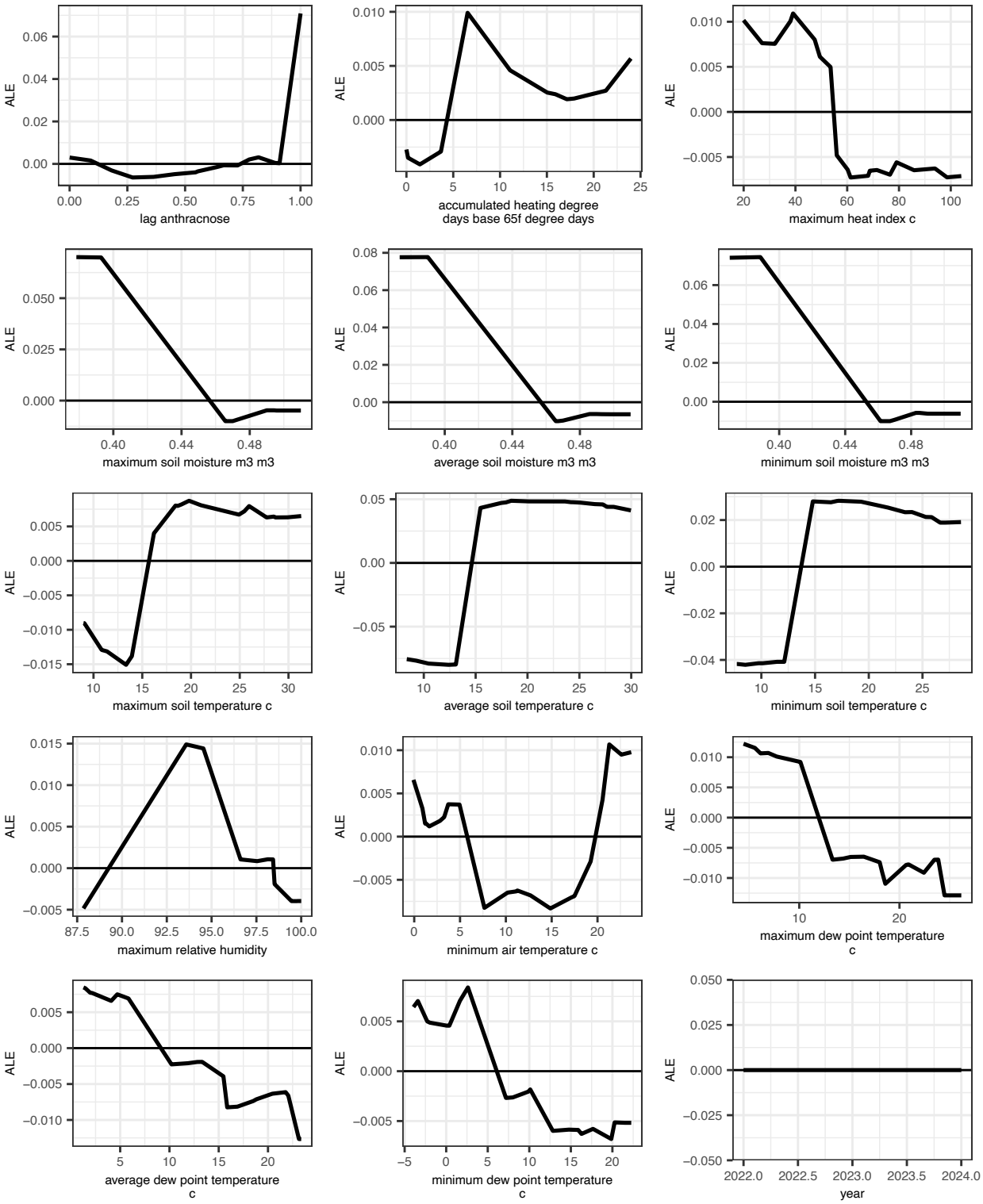

Figure S13. ALE plots for top 15 most important variables in anthracnose random forest model.

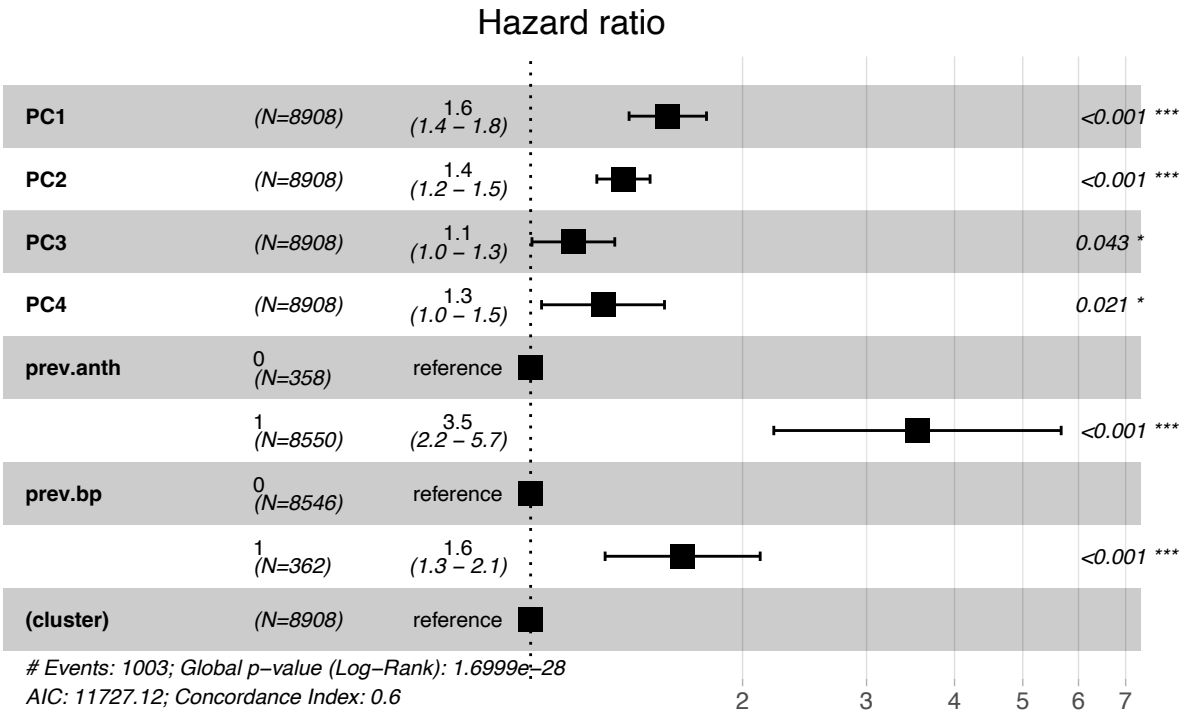

Figure S14. Cox proportional hazard model results for crown rust infection. All environmental PCs significantly increase the odds of observing an anthracnose infection, however prior infection by anthracnose has a stronger effect on the subsequent risk of rust infection.

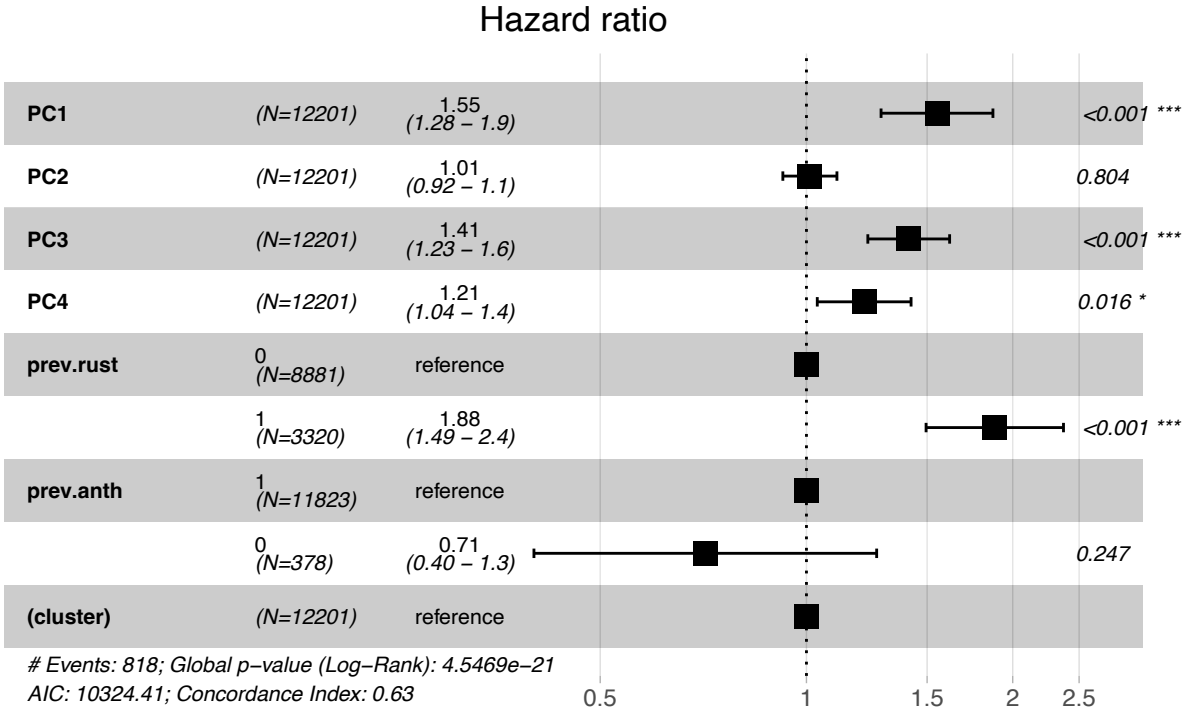

Figure S15. Cox proportional hazard model results for brown patch infection. Environmental PC1, PC3, and PC4 significantly increase the risk of brown patch infection on a similar scale as prior rust infection. Prior anthracnose infection does not affect the risk of subsequent Brown patch infection.

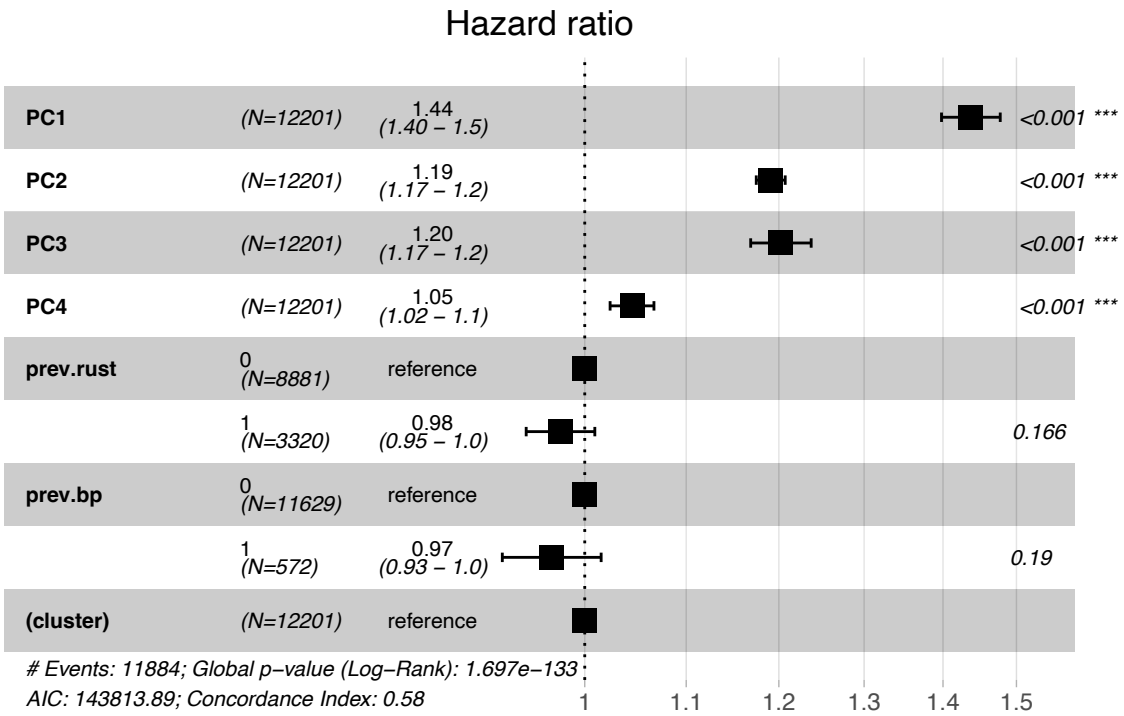

Figure S16. Cox proportional hazard model results for anthracnose infection. All environmental PCs significantly increase the odds of observing an anthracnose infection, whereas prior infection by rust or brown patch do not affect the subsequent odds of observing an anthracnose infection.

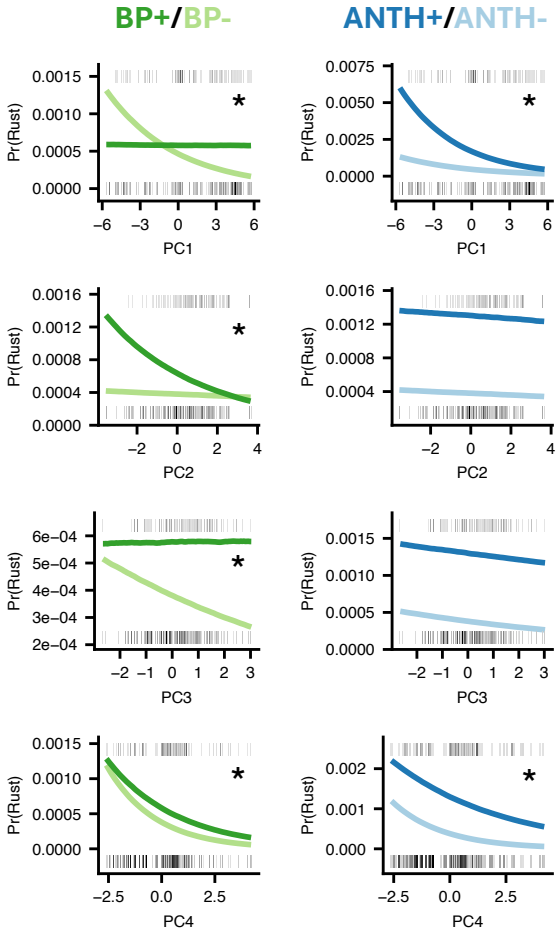

Figure S17. Conditional effects of environmental PCs and coinfection status on the probability of observing crown rust infection from Bayesian hierarchical models. Rugs along the top and bottom of each plot give the observed presences (top) and absences (bottom) of crown rust in the data across PC scores. Plots marked with (*) indicate interaction terms for which zero was not included in the 95% credibility interval of the posterior distribution. (Left) Environmental PC x brown patch conditional effect plots, where the darker green indicates brown patch coinfection. (Right) Environmental PC x anthracnose conditional effect plots, where the darker blue indicates anthracnose coinfection.

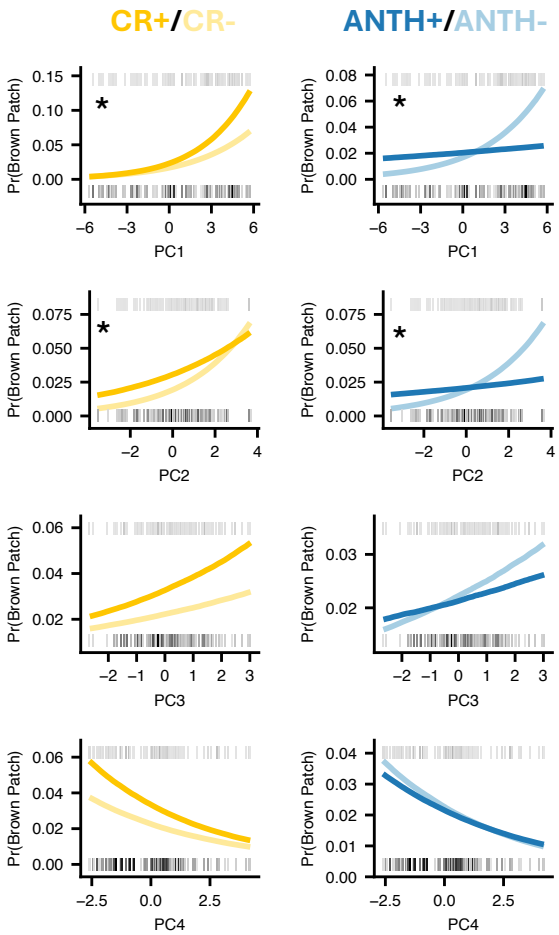

Figure S18. Conditional effects of environmental PCs and coinfection status on the probability of observing brown patch infection from Bayesian hierarchical models. Rugs along the top and bottom of each plot give the observed presences (top) and absences (bottom) of brown patch in the data across PC scores. Plots marked with (*) indicate interaction terms for which zero was not included in the 95% credibility interval of the posterior distribution. (Left) Environmental PC x crown rust conditional effect plots, where the darker yellow indicates crown rust coinfection. (Right) Environmental PC x anthracnose conditional effect plots, where the darker blue indicates anthracnose coinfection.

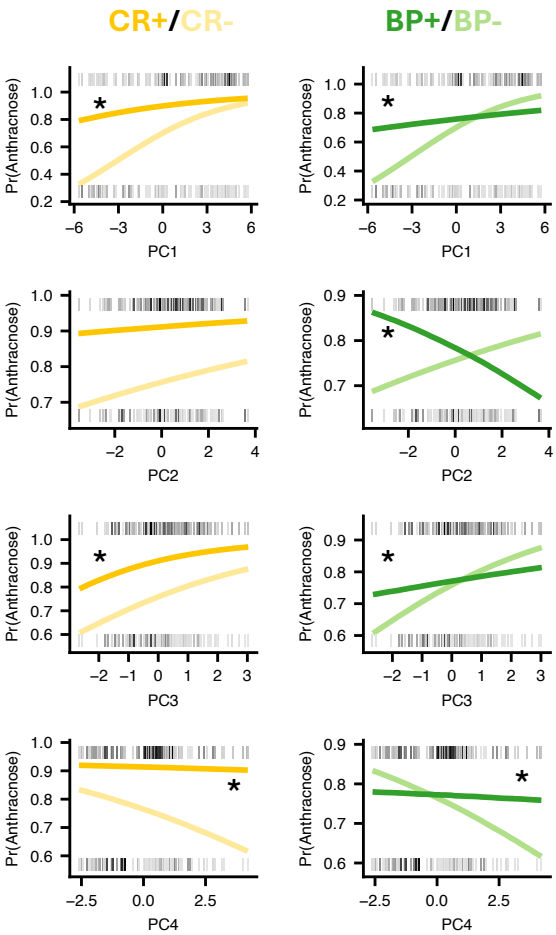

Figure S19. Conditional effects of environmental PCs and coinfection status on the probability of observing anthracnose infection from Bayesian hierarchical models. Rugs along the top and bottom of each plot give the observed presences (top) and absences (bottom) of anthracnose in the data across PC scores. Plots marked with (*) indicate interaction terms for which zero was not included in the 95% credibility interval of the posterior distribution. (Left) Environmental PC x crown rust conditional effect plots, where the darker yellow indicates crown rust coinfection. (Right) Environmental PC x brown patch conditional effect plots, where the darker green indicates brown patch coinfection.

Supplemental Tables

Table S1. Variables included in random forest models from DURH weather station.

Variables Retrieved from DURH	**Unit**
Average Wind Direction	degree
Minimum wind speed	kph
Average Wind Speed	kph
Maximum Wind Speed	kph
Average Relative Humidity	%
Minimum Relative Humidity	%
Maximum Experimental Leaf Wetness	mV
Total Precipitation	in
Maximum Relative Humidity	%
Minimum Air Temperature	C
Minimum Dew Point Temperature	C
Maximum Dew Point Temperature	C
Average Dew Point Temperature	C
Maximum Soil Temperature	C
Minimum Soil Temperature	C
Growing Degree Days per Day - Base 50F	Degree Days
Average Air Temperature	C
Maximum Heat Index	C
Maximum Air Temperature	C
Maximum Soil Moisture	m3/m3
Average Soil Moisture	m3/m3
Minimum Soil Moisture	m3/m3
Average Solar Radiation	W/m2
Average Photosynthetically Active Radiation	W/m2
Maximum Solar Radiation	W/m2
Estimated Evapotranspiration (Penman-Monteith)	cm
Maximum Photosynthetically Active Radiation	micromol/m2s
Average Wind Vector - v Component	kph
Average Wind Vector - u Component	kph
Accumulated heating Degree Day (Base 65F)	degree days

Table S2. (a) **Model performance parameters for crown rust, anthracnose, and brown patch models with 7-day average of environmental variables leading up to sampling date. (b) Descriptions of model metrics used to assess model performance.**

(a)

Metric	*Anthracnose*	*Crown Rust*	*Brown Patch*
Accuracy	0.546	0.808	0.844
Precision	0.529	0.904	0.944
Recall	0.949	0.869	0.883
Kappa	0.0803	0.285	0.197
F-Score	0.679	0.886	0.913

(b)

Metric	*Description*
Accuracy	The proportion of correct predictions made by the model. A value of 1 means the model always predicts correctly, while 0 means it always predicts incorrectly.
Precision	The accuracy of positive predictions, where high precision, approaching 1, indicates few false positives, while low precision, near zero, indicates many false positives.
Recall	How often the model predicts false negatives, where high recall (1) indicates the model did not classify any false negatives, while low recall indicates many false negatives (0).
Kappa	Measures how well the model predicts outcomes compared to random chance. A value of 0 indicates performance no better than chance, and 1 indicates perfect prediction.
F-Score	The harmonic mean of precision (the ability to avoid false positives) and recall (the ability to avoid false negatives). Higher F-Scores, closer to 1, indicate a balance of high precision and recall. This metric is particularly useful when the number of true positive cases is low, as is the case for many of our pathogens.

Table S3. Random forest model metrics using the 14-day average of environmental variables leading up to the sampling date.

Pathogen	Accuracy	Precision	Recall	Kappa	F-Score
Anthracnose	0.535	0.524	0.888	0.0597	0.659
Crown Rust	0.855	0.912	0.921	0.393	0.916
Brown Patch	0.809	0.964	0.824	0.260	0.888

Table S4. Random forest model metrics using the 30-day average of environmental variables leading up to the sampling date.

Pathogen	Accuracy	Precision	Recall	Kappa	F-Score
Anthracnose	0.545	0.532	0.856	0.0817	0.656
Crown Rust	0.810	0.902	0.873	0.278	0.887
Brown Patch	0.792	0.965	0.804	0.242	0.877

Table S5. **Direction of environmental PC effects and their interaction with coinfecting pathogens on focal pathogen presence from Bayesian hierarchical model.** Columns give the focal pathogen (italicized), which were the response variable for each model. Rows give the base effect of each environmental PC on that focal pathogen, and the direction of the interaction between each environmental PC and the presence of another coinfecting pathogen (**bolded**). Direction is reported only for fixed effects with a posterior probability greater than 95% of being non-zero. Fixed effects where this was not the case are indicated by n (no effect).

Enviro PC	**PC1**			**PC2**			**PC3**			**PC4**		
Focal Patho	*ANTH*	*CR*	*BP*	*ANTH*	*CR*	*BP*	*ANTH*	*CR*	*BP*	*ANTH*	*CR*	*BP*
Base Effect	+	-	+	+	*n*	+	+	-	+	-	-	-
ANTH+		+	-		*n*	-		*n*	*n*		+	*n*
CR+	-		+	*n*		-	+		*n*	+		*n*
BP+	-	-		-	-		-	+		+	+	

Table S6. Full Bayesian hierarchical model results for anthracnose.

Variable	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	-187	40.7	-266	-108	1	2160	2670
lag.anthracnose	0.965	0.0724	0.824	1.11	1	2650	2740
month2	0.251	0.117	0.024	0.487	1	1960	2410
month3	-1.08	0.113	-1.3	-0.85	1	1330	2300
month4	0.537	0.128	0.279	0.785	1	1410	2430
month5	0.183	0.184	-0.176	0.544	1	1150	1950
month6	1.42	0.21	1.01	1.82	1	1240	1830
month7	-0.0143	0.208	-0.427	0.395	1	1190	2180
month8	-0.0042	0.213	-0.426	0.409	1	1110	2010
month9	-0.422	0.189	-0.791	-0.05	1	1180	1940
month10	0.13	0.155	-0.184	0.428	1	1230	2230
month11	-0.113	0.111	-0.332	0.101	1	1450	2470
month12	-0.656	0.0983	-0.845	-0.46	1	2180	2890
year	0.0924	0.0201	0.0536	0.132	1	2170	2680
PC1	0.277	0.0199	0.238	0.315	1	1360	1900
PC2	0.0966	0.0188	0.0591	0.134	1	2610	2580
PC3	0.269	0.0258	0.219	0.319	1	2280	2840
PC4	-0.169	0.0284	-0.224	-0.11	1	2200	2720
brown_patch1	0.425	0.112	0.21	0.644	1	2870	2910
crown_rust1	1.33	0.0828	1.18	1.5	1	2940	2990
brown_patch1:crown_rust1	-0.755	0.133	-1.01	-0.5	1	2840	2990
PC1:brown_patch1	-0.213	0.0246	-0.261	-0.17	1	2960	2950
PC2:brown_patch1	-0.251	0.0512	-0.354	-0.15	1	2740	2590
PC3:brown_patch1	-0.184	0.0628	-0.303	-0.05	1	2910	2860
PC4:brown_patch1	0.15	0.0463	0.0593	0.24	1	2700	2810
PC1:crown_rust1	-0.126	0.0209	-0.166	-0.09	1	2990	2880
PC2:crown_rust1	-0.0344	0.0485	-0.13	0.062	1	3160	2890
PC3:crown_rust1	0.102	0.0516	-4E-04	0.204	1	2730	2850
PC4:crown_rust1	0.136	0.0403	0.0586	0.215	1	2720	2660

Table S7. Full Bayesian hierarchical model results for crown rust.

Variable	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	388	40.7	308	468	1	2880	2850
lag.crown_rust	2.39	0.0813	2.23	2.55	1	2800	2750
month2	-0.26	1	-2.44	1.45	1	2950	2660
month3	-1.89	1.3	-4.72	0.282	1	2810	2830
month4	-0.951	0.913	-2.92	0.656	1	2920	2720
month5	-0.275	0.984	-2.43	1.48	1	2640	2500
month6	2.63	0.391	1.87	3.41	0.999	2560	2580
month7	4.79	0.375	4.08	5.54	1	2560	2840
month8	5.15	0.37	4.45	5.89	1	2560	2590
month9	6.5	0.361	5.8	7.24	1	2550	2690
month10	6.28	0.348	5.62	6.98	1	2510	2560
month11	4.29	0.338	3.67	4.98	1	2490	2530
month12	1.96	0.361	1.29	2.72	0.999	2490	2690
year	-0.196	0.0201	-0.235	-0.156	1	2880	2880
PC1	-0.184	0.027	-0.235	-0.132	1	3030	3120
PC2	-0.0276	0.0448	-0.118	0.0602	1	2990	2910
PC3	-0.116	0.0604	-0.236	0.000232	1	3010	2830
PC4	-0.442	0.0457	-0.531	-0.354	1	2520	2770
anthracnose1	1.27	0.0832	1.11	1.43	1	2780	2830
brown_patch1	0.289	0.163	-0.0283	0.606	1	3050	2950
anthracnose1:brown_patch1	-0.708	0.135	-0.967	-0.443	1	3000	2790
PC1:anthracnose1	-0.0408	0.0191	-0.0771	-0.00232	1	3110	2640
PC2:anthracnose1	0.0143	0.0463	-0.0747	0.106	1	3010	2710
PC3:anthracnose1	0.0805	0.0623	-0.0394	0.207	1	2990	2730
PC4:anthracnose1	0.237	0.0417	0.155	0.322	1	2770	2820
PC1:brown_patch1	0.181	0.0257	0.131	0.231	1	2840	2800
PC2:brown_patch1	-0.182	0.0528	-0.288	-0.0805	1	2790	2920
PC3:brown_patch1	0.119	0.0623	-0.00178	0.24	0.999	3190	2700
PC4:brown_patch1	0.137	0.0477	0.0419	0.226	1	3000	2900

Table S8. Full Bayesian hierarchical model results for brown patch.

Variable	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	920	48.4	827	1010	1	2540	2360
lag.brown_patch	0.801	0.122	0.563	1.05	1	3050	2950
month2	0.155	0.394	-0.628	0.901	1	1840	2370
month3	-0.0477	0.411	-0.86	0.747	1	1830	2430
month4	0.0934	0.344	-0.59	0.77	1	1610	2120
month5	0.0448	0.391	-0.713	0.813	1	1480	2290
month6	0.213	0.373	-0.516	0.957	1	1420	2180
month7	1.32	0.38	0.584	2.07	1	1460	2080
month8	2.02	0.373	1.29	2.77	1	1430	2120
month9	1.75	0.355	1.06	2.45	1	1400	1970
month10	1.84	0.322	1.22	2.51	1	1410	2010
month11	0.594	0.305	0.0319	1.21	1	1510	2370
month12	0.221	0.343	-0.444	0.914	1	1800	2270
year	-0.458	0.0239	-0.504	-0.412	1	2550	2360
PC1	0.258	0.0352	0.188	0.327	1	2410	2620
PC2	0.362	0.0475	0.266	0.456	1	2530	2850
PC3	0.125	0.0655	-0.00486	0.256	1	2490	2770
PC4	-0.203	0.0498	-0.303	-0.107	1	2780	2630
anthracnose1	0.328	0.107	0.116	0.538	1	2490	2780
crown_rust1	0.398	0.15	0.101	0.691	1	2550	2460
anthracnose1:crown_rust1	-0.576	0.135	-0.834	-0.304	1	2680	2850
PC1:anthracnose1	-0.215	0.0235	-0.259	-0.169	1	2710	2600
PC2:anthracnose1	-0.284	0.0499	-0.382	-0.187	1	2680	2660
PC3:anthracnose1	-0.0541	0.0676	-0.188	0.075	1	2490	2770
PC4:anthracnose1	0.029	0.0483	-0.0638	0.123	1	2890	2990
PC1:crown_rust1	0.0629	0.0245	0.0151	0.112	1	2700	2780
PC2:crown_rust1	-0.163	0.0474	-0.256	-0.07	1	2780	2760
PC3:crown_rust1	0.0428	0.0542	-0.062	0.151	1	2850	2800
PC4:crown_rust1	-0.0182	0.0445	-0.108	0.0685	1	2950	2910