BIODIVERSITY RESEARCH

Diversity and Distributions   WILEY

The distribution of a group of keystone species is not associated with anthropogenic habitat disturbance

Benjamin R. Fitzpatrick[1] | Andri Baltensweiler[2] | Christoph Düggelin[2] |
Marielle Fraefel[2] | Anne Freitag[3,4] | Martijn L. Vandegehuchte[1,5] |
Beat Wermelinger[6] | Anita C. Risch[1]

[1]Community Ecology Research Unit, Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland

[2]Forest Resources and Management Research Unit, Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland

[3]Cantonal Museum of Zoology, Lausanne, Switzerland

[4]Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

[5]Terrestrial Ecology Unit, Department of Biology, Ghent University, Ghent, Belgium

[6]Forest Health and Biotic Interactions Research Unit, Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland

Correspondence
Benjamin R. Fitzpatrick, Community Ecology Research Unit, Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland.
Email: benjamin.fitzpatrick@wsl.ch

Abstract

Aim: Red wood ants (*Formica rufa* group) mitigate invertebrate pest outbreaks, alter invertebrate communities, and contribute to nutrient cycling. The IUCN lists these insects as near threatened. We investigated whether disturbances of Red Wood Ant (RWA) forest habitats related to recreation, infrastructure, and forest management were associated with RWA occurrence. We also investigated the habitat associations of this group.

Location: Switzerland.

Methods: We trained random forest models to predict RWA occurrence using data from 6,341 plots distributed throughout Swiss forests. Our explanatory variables included descriptions of vegetation, terrain, climate and human disturbance of the forest. A model trained using all of these variables (Model 1) was compared to another trained using all of these variables except those describing human disturbance (Model 2). We compared the abilities of these models to differentiate between RWA presences and absences using areas under receiver operator curves (AUC). The nature of the associations between the probability of RWA occurrence predicted by Model 1 and the explanatory variables that made the greatest contributions to the AUC of this model were investigated using individual conditional expectation plots.

Results: No significant difference in AUC was detected between Models 1 and 2. RWA occurrence was positively associated with elevation, conifers, canopy gaps, ground cover vegetation and solar insolation while negatively associated with air temperature and a soil wetness index.

Main conclusions: The distribution of RWA within Swiss forests appeared unassociated with the human disturbances we investigated. To conserve RWA in Switzerland, we recommend the conservation of forests with high proportions of conifers, particularly those at high elevations. We also recommend forest management that promotes ground cover vegetation and open canopy structures. The negative associations between RWA occurrence and temperatures raise concerns for the prospects of RWA in the context of predicted climate warming.

1 | INTRODUCTION

Reductions in species interactions accompany the human-induced decline in biodiversity and may reduce the capacity of ecosystems to provide functions and services (Cardinale et al., 2012; Pimm et al., 2014; Tilman et al., 2014; Valiente-Banuet et al., 2015). Keystone species and ecosystem engineers interact with disproportionately large numbers of other species and are thus worthy of inclusion in conservation strategies. Red wood ants (*Formica rufa* group) occur throughout the forests of the northern hemisphere (Stockan et al., 2016) and have been identified as keystone species (Stockan et al., 2016) and ecosystem engineers (Sorvari, 2016). Red wood ants (hereafter RWA): have mutualistic relationships with phloem-feeding insects (Domisch et al., 2016); protect host trees from certain invertebrate pests (Laine & Niemela, 1980; Maňák et al., 2013); are prey for several vertebrate species (Robinson et al., 2016); and substantially alter the invertebrate communities in the vicinity of their mounds (Robinson et al., 2016). In addition, RWA mounds provide habitats for a variety of species (Robinson et al., 2016). RWA also influence soil properties and ecosystem functions in and around their mounds (Jurgensen et al., 2008).

Recognition of the positive roles RWA can play in forest ecosystems and concern about their distribution, abundance and prospects have motivated several European countries to create laws protecting RWA and their mounds (Sorvari, 2016). As reviewed by Sorvari (2016), threats to RWA include both habitat destruction (e.g. conversion of forest to other land uses), and changes in climate and microclimate. The RWA species *F. aquilonia* Yarrow, *F. lugubris* Zettersted, *F. polyctena* Förster, *F. pratensis* Retzius, *F. rufa* Linnaeus and *F. uralensis* Rutzsky have been included on the IUCN Red List as near threatened species since the 1960s, though data on these species remains insufficient to update the original listings (IUCN, 2019). The absence of other RWA species from the IUCN Red List reflects a lack of knowledge rather than confidence in the prospects of these species (Sorvari, 2016).

Conservation of RWA can be aided by understanding the major influences on the distribution of these insects. However, to our knowledge, only three such studies have been conducted using data collected across areas of 1,000 km^2 or greater. In the United Kingdom, Procter et al. (2015) detected a gradual expansions of *F. lugubris* from the edges of older forest into more recently reforested areas. Using data describing the surroundings of RWA mounds and pseudo-absences, Procter et al. (2015) detected a positive association between *F. lugubris* occupancy and the proportion of conifers in these surroundings and a negative association between *F. lugubris* occupancy and the proportion of open ground in these surroundings. Procter et al. (2015) concluded that plantations of non-native conifers can provide suitable habitat for *F. lugubris* and

recommended that such plantations be established in areas adjoining currently forested areas to create new habitat into which this species may expand. Dekoninck et al. (2010) attributed extinctions of RWA colonies (*F. rufa* and *F. polyctena*) in their study area in Belgium to increased shading by vegetation, recreational use of forests and pollution from nearby agriculture. Dekoninck et al. (2010) recommended the conservation of existing RWA habitat as a first priority along with: increasing connectivity between patches of existing habitat; promoting varied vegetation structures and gaps in the forest canopy; increasing the amount of sunlight that reaches the soil at forest edges; and creating buffer zones between RWA colonies and agricultural land. Vandegehuchte et al. (2017) found sixteen explanatory variables that described vegetation, climate, aspect and insolation to be useful predictors of RWA occurrence (mainly *F. lugubris* and *F. paralugubris*) in Switzerland. Of these predictors, canopy cover by conifers and climate continentality were assigned the two largest magnitude coefficients in their model, followed by eastness (aspect) and solar radiation in the month of July, all four of which had positive associations with RWA occurrence. To conserve RWA, Vandegehuchte et al. (2017) recommended the conservation or establishment of coniferous forests with diverse structures, open canopies and a high degree of cover by ground vegetation. Of these three broad-scale studies, only Dekoninck et al. (2010) considered some estimates of direct human impacts on the environment surrounding the ant mounds. This study and others conducted over smaller spatial scales (see below) provide us with material from which to hypothesise about the potential influences human-related disturbances may have on RWA.

Forest management in its most destructive form of clearcutting has detrimental effects on RWAs (Sorvari, 2016), similar to those arising from the conversion of forest to other land uses. However, clear cutting of Swiss forests is prohibited except under special circumstances (Federal Forest Act, 1991). The environmental disturbances accompanying urbanisation have been attributed as the cause of decreased densities of RWA mounds in Espoo, Finland (Vepsäläinen & Wuorenrinne, 1978; Wuorenrinne, 1989). However, both of these studies lacked any statistical testing of this hypothesis. Roadside verges in forested areas may provide suitable habitat for some RWA species by functioning as artificial forest edges. However, Freitag et al. (2008) observed that frequent mowing, pollution, and low supplies of food decrease the suitability of roadside verges as habitat for *F. pratensis* in south western Switzerland. In contrast, Perron et al. (2019) claim to have detected positive associations between densities of *F. rufa* and *F. polyctena* and proximities to roads or paths independent of the bias in their data introduced by primarily conducting their sampling from these roads and paths. The impacts of recreational use of forests on RWA colonies have not been extensively studied to our knowledge. Bugrova and Reznikova

(1990) described increasing percentages of monodomous colonies from the typically polydomous *F. polyctena* accompanying increasing trampling of ground vegetation from increasing recreational pressure in forest near an urban centre in Siberia. However, they also described *F. polyctena* colonies occurring in plots where up to 50% of the ground cover vegetation had been trampled. Furthermore, they attempted no statistical analysis of the differences they reported between plots with differing amounts of recreational pressure. Dekoninck et al. (2010) documented the loss of thirty wood ant colonies in North Western Belgium and attributed the loss of seven of these colonies to recreational pressure.

In this study, we have expanded upon previous modelling studies of the factors influencing RWA distributions across areas of 1,000 km^2 or greater by including among the explanatory variables descriptions of human disturbance of the forest. We included data on: forest management activities which involved disturbances of the forest less severe than complete habitat destruction; recreational use of the forest; and proximity to human structures such as roads and buildings in addition to the more commonly used descriptions of climate, terrain, and vegetation. We compared the predictive performance of a model fitted using all of our explanatory variables (Model 1) to that of a model fitted using all of our explanatory variables except for those that quantified direct human disturbance of the forest (Model 2). We investigated the nature of the associations between the most useful explanatory variables and the probability of RWA occurrence predicted by Model 1 and searched for evidence of interactions between these variables. Thus, in addition to considering a more comprehensive set of explanatory variables than any previous broad-scale study of RWA habitat associations, we are the first to examine potential interactions between explanatory variables. Furthermore, following Procter et al. (2015), our study is the second to consider the potential for non-linear effects of explanatory variables in a broad-scale model of RWA distributions. We hypothesised that habitat disturbance by humans would negatively affect RWA. Specifically, we predicted that: more recent and destructive forest management; higher recreational use by humans; and shorter distances to buildings would be negatively associated with RWA occurrence. In contrast, we predicted that proximity to roads would be positively associated with RWA occurrence. Overall, we predicted that Model 1 would outperform Model 2.

2 | METHODS

2.1 | Data sources and preparation

We used data collected by the fourth Swiss National Forest Inventory (hereafter NFI) which was conducted between 2009 and 2017. The NFI collected field data from 12.62 m radius (500 m^2) plots on the vertices of a 1.4 km grid that covered all forested areas of Switzerland. In addition, an interview survey was conducted with local forestry services concerning the forest management history and recreational use of the forest management units that contained each NFI plot. The NFI protocols are detailed in the NFI field manual (Düggelin & Keller, 2017) and the NFI interview survey manual (Keller, 2013). The fourth NFI was the first NFI in which each plot was searched for RWA mounds. When an ant mound was detected, approximately 10 ants were collected for species identification. The species identification process is detailed in Vandegehuchte et al. (2017). With these data, we were able to assert the presence or absence of RWA mounds in 6,341 plots (see Appendix S1 in the Supporting Information). There was a substantial class imbalance in the RWA occurrence data with RWA mounds present in only 290 (4.57%) of these 6,341 plots.

We prepared data from the NFI concerning vegetation and human disturbance of the forest for use in our models. These data included descriptions of the structure, composition, age and origin of vegetation within plots. We also included the presence or absence of standing or flowing water in plots along with descriptions of evidence within plots of snowpack movement, avalanche, windthrows, and other damage to the ground and vegetation. In addition, we included estimates of lengths of time that had elapsed between the most recent damage and the date when the plot was searched for RWA mounds. Other NFI data we included described: forest usage; forestry intervention types; the time elapsed since the most recent forestry intervention; and transport access to plots. Our data on human disturbance of forests describe disturbances less severe than the complete habitat destruction that would accompany clearcutting or the conversion of forest to alternative land uses. This is a consequence of: NFI plots only being located within areas categorised as forest (Düggelin & Keller, 2017); the prohibition of clear cutting (Federal Forest Act, 1991); and the legal protection of both the area and spatial distribution of Swiss forests (Federal Forest Act, 1991). We also included descriptions of evidence of grazing and the presence of structures in plots along with estimates of the types and intensities of recreational uses of the forest. The full set of NFI variables we selected and our preparation of these variables for use in our models is presented in Appendices S2 and S3 in the Supporting Information.

We prepared a collection of terrain attributes (TA) from a digital elevation model (DEM) with a 5 m resolution that had been resampled from a DEM with a 2m resolution (SwissAlti3D). We calculated the Cell Balance, Flow Accumulation, Mass Balance Index, Melton Ruggedness Number, Site Exposure, Slope, Topographic Wetness Index, and Valley Depth with SAGA (Conrad et al., 2015). We calculated the plan and profile curvature of the terrain with ArcGIS (ESRI, 2010a). Using the Python Framework of ArcGIS 10.5, we calculated the following TA: Topographic Position Index (Weiss, 2001), Heatload (McCune & Keon, 2002), Integrated Moisture Index (Iverson et al., 1997), Martonne's Modified Dissection (Evans, 1972), Roughness, Eastness, Northness, and Global Potential Radiation (Fu & Rich, 2002). We described the distributions of elevation and TA pixel values within each NFI plot using sets of summary statistics. For each NFI plot and each variable associated with DEM pixels, we calculated the minimum, first quartile, mean, median, third quartile, maximum, standard deviation, and Moran's I of the values of that

variable associated with the DEM pixels which intersected the NFI plot (see Appendix S4 in the Supporting Information).

We obtained daily climate interpolations from 1999 to 2016 at a 100 m resolution across all of Switzerland. These were produced by applying the Daymet method (Thornton et al., 1997) to daily climate observations from Swiss weather stations (MeteoSwiss) and a 100 m DEM (SwissTopo). These data contained daily values for: precipitation; average, minimum and maximum temperature; insolation; and potential evapotranspiration. We calculated 19 BIOCLIM (Busby, 1991; Nix, 1986) and 16 ANUCLIM (Xu & Hutchinson, 2011) style statistics from these data. We also calculated BIOCLIM style interval-based statistics for our estimates of the active and dormant periods of RWA. We calculated the above set of climate summary statistics at each plot for each of the 10 years that immediately preceded the year in which the plot was searched for RWA mounds. We then reduced these sets of annual values to 10-year means of each summary statistic at each plot (see Appendix S5 in the Supporting Information). We focused on descriptions of the climate over periods of 10 years because the establishment, persistence, and growth of a RWA mound is a process that spans multiple years.

With the topographic landscape model of Switzerland (swissTLM3D) and ArcMap (ESRI, 2010b) we calculated the Euclidean distances from the centre of each plot to the nearest small road, large road and building. This involved grouping the roads in these data into two categories. We defined small roads to range from unsealed trails and 4WD tracks through to sealed country roads of less than 3 m in width while large roads were defined to include all larger and better developed roads, highways, and freeways (see Appendix S6 in the Supporting Information).

2.2 | Dimension reduction

Using the R (R Core Team, 2018) package "ClustOfVar" (Chavent et al., 2012), we identified strongly related groups of explanatory variables and summarised each group with a single synthetic variable. This produced the 129 synthetic explanatory variables we have used in our analysis, reducing both the number of explanatory variables and the degree of collinearity among these variables. Further details on this process are given in Appendix S7 in the Supporting Information.

2.3 | Modelling RWA occurrence

To address our objectives, we sought a method for modelling RWA occurrence that: had the potential to efficiently model linear, non-linear and interacting effects of explanatory variables; was appropriate for use with data that included correlated groups of explanatory variables; and was appropriate for use with a response variable that contained a class imbalance. In logistic regression models, non-linear effects of explanatory variables may be approximated via polynomial terms and interactions between groups of explanatory variables

may be approximated by terms for products of these groups of variables. However, as the number of potential explanatory variables and the complexity of interactions and non-linear effects considered increase, the number of potential terms from which such logistic regression models may be constructed increases combinatorially. Model selection by exhaustive search with such large sets of potential terms can be very computationally demanding. Stepwise variable selection procedures, popular alternatives to exhaustive searches, are best avoided in situations where substantial correlations exist among potential explanatory variables, as is frequently the case with environmental data (Dormann et al., 2013). Furthermore, correlations between explanatory variables cause instability in the maximum likelihood estimates of coefficients in multiple linear regression models (Dormann et al., 2013). The penalized likelihood technique LASSO (Tibshirani, 1996, 2011) efficiently fits parsimonious multiple regression models when supplied with large sets of potential explanatory variables. However, the identities of the variables selected by LASSO from groups of strongly correlated explanatory variables are not meaningful (Zou & Hastie, 2005). In contrast, random forests (Breiman, 2001) are computationally efficient and have the potential to approximate complex non-linear relationships and interactions (Cutler et al., 2007) without additional computational cost. Furthermore, methods have been designed for random forests to rank the relative importance of correlated explanatory variables (see Section 2.5) and address class imbalances within response variables (see below).

We used random forests composed of conditional inference trees (Hothorn, Hornik et al., 2006) as implemented in the R package "party" (Hothorn, Bühlmann et al., 2006) to model RWA occurrence using our 129 synthetic explanatory variables. We used an option provided by the "party" package to substitute a sampling-without-replacement step for the bootstrap sampling step involved in fitting a random forest. This substitution mitigated the bias in variable selection that traditional random forests would have otherwise displayed towards explanatory variables with many unique values (Strobl et al., 2007). This substitution also made the variable importance scores we calculated appropriate for ranking the relative importance of explanatory variables that contained varying numbers of unique values (Strobl et al., 2007) as were present among our data.

We also made use of other features of the "party" package to adjust the model fitting procedure to be appropriate for data in which the response variable displayed a substantial class imbalance. We used the weights option of the "cforest" command to ensure that the expected number of in-bag RWA mound presences was equal to the expected number of in-bag RWA mound absences. Each tree of a random forest is fitted to a set of in-bag observations. Consequently, equal expected numbers of presences and absences in-bag ensure that presences and absences have equal influence on the model fitting process. The replacement of bootstrapping with sampling without replacement in the random forest fitting process (see above) necessitated consideration of the size of the samples drawn without replacement to constitute the in-bag samples. In particular, due to the trees being fitted to

the in-bag observations and the variable importance scores being calculated using the out-of-bag observations it was necessary to ensure that both the in-bag and out-of-bag samples contained presences. We chose an in-bag sample size and sampling weights that combined lead us to expect half of the presences to be in-bag and half of the presences to be out-of-bag in any sample of points drawn prior to fitting a tree of the random forest. The relatively small size of the in-bag samples this required was one of our motivations for seeking to optimize the number of trees in the random forest by cross-validation and considering larger numbers of trees than the common choice of 500 trees (see Section 2.4).

2.4 | Cross-validation informed by spatial dependence structure

We sought to optimise two tuning parameters and the number of trees in our random forests via cross-validation. The first of these tuning parameters was the number of explanatory variables selected at random from the full set each time a partition in a tree was defined. The second of these tuning parameters functions as a stopping criterion in the tree growing process. This parameter is the critical p-value for the tests of association between explanatory variables and the response conducted when seeking to partition a node. If none of the tests return a p-value below this critical value, no further partitions of the node are attempted.

We examined the RWA occurrence (response) data for spatial dependence and used this information when partitioning our data for cross-validation (Roberts et al., 2017). We applied join count tests (Bivand et al., 2013) to regions of different densities of presences to determine the maximum range of spatial dependence among response observations throughout our study area. We used the following three neighbourhood definitions: (a) binary neighbours defined with a distance threshold; (b) inverse distance-based neighbourhood weights; and (c) inverse distance squared-based neighbourhood weights. Having identified a distance beyond which we no longer detected spatial dependence between response observations anywhere in the study area, we defined subsets of the data to serve as validation sets. We mitigated any spatial dependence between observations in associated pairs of training and validation sets by establishing spatial exclusion buffers around the regions that defined validation sets. Any observations that intersected such an exclusion buffer were excluded from both the training and validation sets for that particular partitioning of the data. We set these exclusion buffers to be slightly wider than the maximum range of spatial dependence detected. We defined validation sets using convex regions in space to minimise the loss of observations to exclusion buffers. We defined these regions to contain approximately 20% of the data so that the data outside the exclusion buffers that encompassed these regions could serve as training sets containing approximately 80% of the data. In addition, we defined each validation set to contain a ratio of presences to absences that was similar to that in the full data set. In this manner, the class imbalance in the full data set was

represented in both the training sets and the validation sets. For further detail on the process described above see Appendix S8 in the Supporting Information. We selected tuning parameter values on the basis of the areas under receiver operator curves (AUC values) associated with predicting the response observations from the validation sets in the cross-validation procedure. Our choice of AUC for this purpose was motivated by the class imbalance in our response variable. AUC was suitable for this purpose because it measures the accuracy with which a model can differentiate between presences and absences where other estimates of prediction accuracy would favour models that more accurately predicted the majority class.

2.5 | Model interpretation: Variable importance, model comparisons and ICE plots

We used the "party" package to calculate variable importance scores for the explanatory variables in the random forests we fitted. We used two features of the "party" package when calculating these scores. The first feature enabled us to calculate conditional variable importance (hereafter CVI) scores suitable for determining the relative importance of explanatory variables among which correlations exist (Strobl et al., 2008), as was the case with some of our synthetic explanatory variables. The second feature enabled us to base these CVI scores on AUC (Janitza et al., 2013) which was appropriate in the context of the class imbalance we expected among the out-of-bag samples (see Section 2.3). We examined the stability of the CVI scores by refitting random forests to our data 50 times using the tuning parameter values selected by cross-validation and setting a different random seed each time. We calculated CVI scores for each of these random forests and ranked our explanatory variables in order of decreasing mean CVI. While CVI scores provide a useful ranking of explanatory variable importance, we are not aware of a method for testing the null hypothesis that the CVI score of an explanatory variable is equal to zero. Consequently, we tested our hypothesis regarding the association between human disturbance of the forest and RWA occurrence by comparing the AUC of a model fitted using all 129 of our explanatory variables (Model 1) to the AUC of a model fitted using all of our explanatory variables except for the twelve variables that quantified direct human disturbance of the forest (Model 2). We used the permutational test provided by the R package pROC (Robin et al., 2011) to compare the AUC values of the different models. We then sought a more parsimonious model and fitted a random forest (Model 3) using only the six explanatory variables that had mean CVI scores from Model 1 greater than 10% of the maximum mean CVI score from this model. A random forest (Model 4) that used only the four explanatory variables with the highest mean CVI scores from Model 1 had a very similar AUC value to that of Model 3. Further reductions in the set of explanatory variables used to fit random forests resulted in more substantial decreases in the AUC associated with these simpler models. We compared Models 2, 3 and 4 to Model 1 via permutational tests of differences in AUC values.

TABLE 1 The 10 explanatory variables with the highest mean conditional variable importance (CVI) scores from Model 1

Variable	Description	Mean CVI (Rescaled)	Std. Dev. CVI (Rescaled)
Clim.Terr.C62	Positively correlated with elevation and negatively correlated with temperatures.	100.00	4.37
NFI.S1.C2	A description of forest composition with low values corresponding to dominance by conifers and high values corresponding to dominance by broadleaf trees.	27.87	1.75
Clim.Terr.C69	Positively correlated with the 10 year mean of the mean temperatures during the driest quarters.	24.37	1.84
NFI.S1.C13	Positively correlated with the presence of gaps in the forest and cover by ground cover vegetation and negatively correlated with canopy cover.	16.84	1.04
Clim.Terr.C64	Negatively correlated with global potential radiation.	14.95	0.94
Clim.Terr.C25	Positively correlated with the integrated moisture index (a soil wetness index)	10.43	0.90
Clim.Terr.C28	Positively correlated with northness (aspect) and negatively correlated with site exposure.	5.06	0.69
Clim.Terr.C68	Positively correlated with the 10 year mean of mean temperatures during the wettest quarters.	4.85	0.67
Clim.Terr.C18	Negatively correlated with eastness (aspect)	3.93	0.52
Clim.Terr.C70	Positively correlated with precipitation.	3.74	0.45

Note: The mean CVI scores have been expressed as percentages of the highest mean CVI score. The standard deviations of the CVI scores have been expressed on the same scale.

We examined the nature of the associations between explanatory variables and the probability of RWA occurrence predicted by Model 1 using individual conditional expectation (ICE) plots (Goldstein et al., 2015). We produced ICE plots for each of the six explanatory variables with the highest mean CVI scores from Model 1. Where grouping was apparent among the curves of an ICE plot, we investigated whether this grouping was associated with another explanatory variable. In this manner, we investigated potential interactions among the explanatory variables. This process is detailed in Appendix S9 in the Supporting Information.

The R code written to perform the data preparation, analysis and interpretation is available from the repository located at https://github.com/brfitzpatrick/rwaha. This repository includes a machine-readable list of all R packages used.

3 | RESULTS

We detected RWA mounds in 290 NFI plots, which corresponds to approximately 4.57 % of the plots for which we had occurrence data. RWA were more common in the Alps and Jura mountains than in the Swiss lowlands (see Appendix S1). The majority of the RWA mounds

detected were inhabited by *F. lugubris* or *F. paralugubris*, though we also detected mounds inhabited by *F. aquilonia*, *F. polyctena*, *F. pratensis* and *F. rufa* (see Appendix S1).

3.1 | Model comparisons

Model 1, which included all of our explanatory variables, had an AUC of 0.89853. Model 2 lacked variables describing direct human disturbance of the forest yet had a very similar AUC (0.89847). The difference between these two AUC values was not significantly different from zero ($p = 0.926$). Thus, the inclusion of descriptions of direct human disturbance of the forest did not improve the predictive accuracy of a model that used only descriptions of climate, vegetation and terrain. Model 3, which included only the six explanatory variables with the highest mean CVI scores from Model 1, had an AUC of 0.868101. Model 4, which used only the four explanatory variables with the highest mean CVI scores from Model 1, had an AUC of 0.8505975. The AUC values of both Model 3 and Model 4 differed significantly from that of Model 1 ($p < 0.0001$ in both cases), however, the magnitudes of these differences in AUC were relatively small.

3.2 | Variable importance and associations with predicted values

Of the variables used in Model 1, Clim.Terr.C62 had the highest mean CVI (Table 1) and was positively associated with RWA occurrence (Figure 1a). Clim.Terr.C62 is positively associated with elevation and negatively associated with temperature. This suggests a higher probability of RWA occurrence at higher elevations, where temperatures are lower. The variable with the second highest mean CVI was NFI.S1.C2, which described the species composition of the trees in a plot. Low values of NFI.S1.C2 corresponded to plots dominated by conifers whereas high values of NFI.S1.C2 corresponded to plots dominated by broadleaf trees. RWA occurrence was negatively associated with the degree to which broadleaf trees dominated a plot (Figure 1b) and consequently positively associated with the degree to which conifers dominated a plot. The variable with the third highest mean CVI, Clim.Terr.C69, represents the 10-year mean of temperatures during the driest quarter of the year and was negatively associated with RWA occurrence (Figure 1c). At most NFI plots the driest quarter of the year occurred at least partially during winter (see Appendix S9). This suggests a trend of RWA either

avoiding, failing to colonize or failing to persist in areas with a period of comparatively warm temperatures during their dormant period. Clim.Terr.C69 also displayed a strong negative correlation with Clim.Terr.C62 ($r = -0.87$), which suggests that periods of warmer temperatures during the dormant period are predominantly a characteristic of lower elevations where the overall mean temperatures are higher. The positive association between NFI.S1.C13, which had the fourth highest mean CVI, and RWA occurrence suggests that gaps in the forest, openness in the upper layer of the forest and the density of ground cover vegetation are important for RWA (Figure 1d). The negative association between Clim.Terr.C64, which had the fifth highest mean CVI, and RWA mound occurrence (Figure 1e) suggests a positive effect of solar insolation on RWA mounds. RWA occurrence was negatively associated Clim.Terr.C25 (Figure 1f) which represented a soil wetness index and had the sixth highest mean CVI. We also found evidence for interactions between Clim.Terr.C62 and the associations of both Clim.Terr.C64 and Clim.Terr.C25 with RWA occurrence (Figure 2). More specifically, we discovered that the two groups of curves in each of the associated ICE plots could be effectively separated using the values of Clim.Terr.C62 associated with each curve. In the case of Clim.Terr.C64, this separation suggests
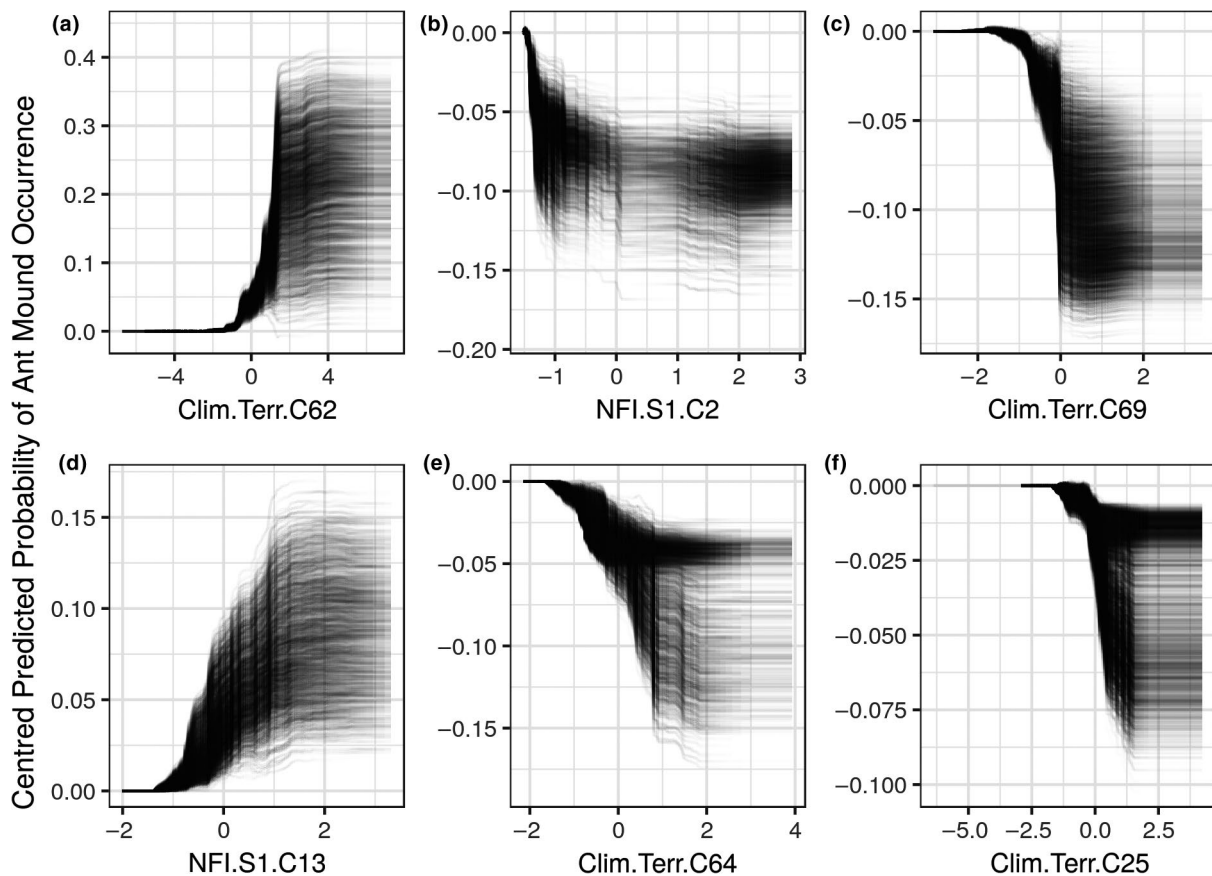

FIGURE 1 Individual conditional expectation plots depicting the nature of the associations between each of the six most important explanatory variables in Model 1 and probability of RWA occurrence predicted by this model. Each line represents predictions obtained by evaluating the model at a sequence of values of the explanatory variable on the horizontal axis of the plot while holding all other explanatory variables constant at the combination of values they took at one of the observations in the data. The opacity of a segment of a line is inversely proportional to the difference between the value of the explanatory variable on the horizontal axis of the plot associated with that line segment and the value that explanatory variable took at the observation associated with the line. Acronyms are defined in Table 1
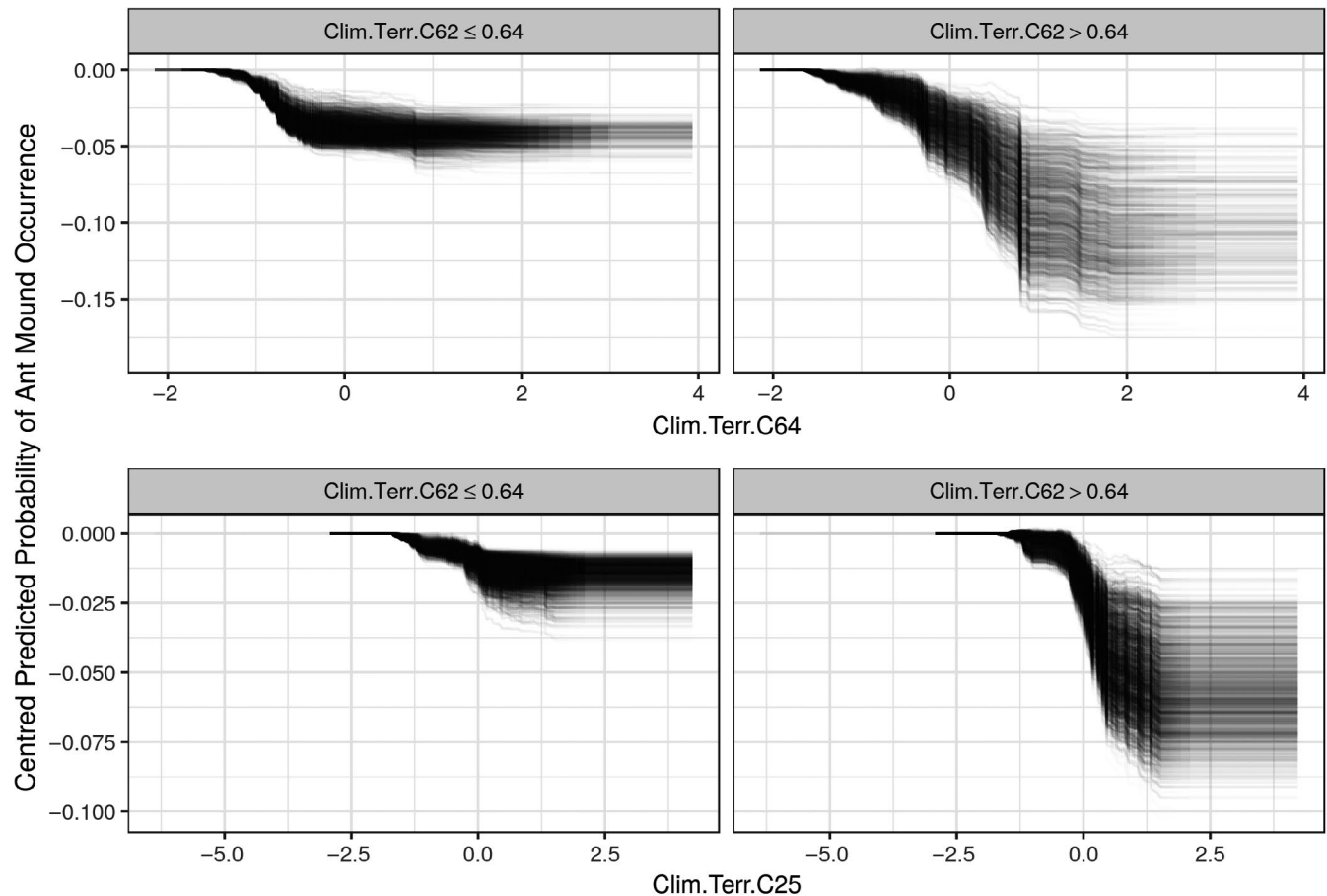
FIGURE 2 Individual conditional expectation plots depicting apparent interactions between Clim.Terr.C62 and each of Clim.Terr.C64 and Clim.Terr.C25. Increases in both of Clim.Terr.C64 and Clim.Terr.C25 are accompanied by decreases in the predicted probability of ant mound occurrence. Higher values of Clim.Terr.C62 were associated with more extreme decreases in predicted probability of ant mound occurrence as each of Clim.Terr.C64 and Clim.Terr.C25 increased. Acronyms defined in Table 1

that insolation was more important for RWA at higher elevations where temperatures were lower (Figure 2). Similarly, this separation of the ICE curves for Clim.Terr.C25 by Clim.Terr.C62 suggests that the property being measured by the soil wetness index had a more extreme negative influence on RWA mound occurrence at higher elevations (Figure 2).

4 | DISCUSSION

4.1 | Anthropogenic habitat disturbances and RWA

The main goal of this paper was to assess whether models that included variables describing human disturbance of the forest would more accurately predict the occurrence of RWA than models fitted to data that lacked these variables. Against our expectations, the variables describing human disturbances we investigated did not significantly improve the accuracy of models for RWA mound occurrence that contained descriptions of vegetation, terrain and climate. It is possible that some or all of these disturbances have affected RWA abundance rather than occurrence. However, negative effects of disturbances that have not resulted in colony extinctions will not

be apparent from the results of our analysis of presence-absence data and reliable abundance data for RWA in Switzerland are not currently available.

Our finding that the forest management-related disturbances we investigated did not appear to have a negative effect on RWA occurrence could be a consequence of forest management in Switzerland having less severe impacts on RWA than those reported in other countries (Sorvari, 2016). For example, clear cutting of Swiss forests is generally prohibited with the exception of special circumstances and both the area and spatial distribution of Swiss forests are protected (Federal Forest Act, 1991). The apparent lack of influence of the recreational activities we investigated on RWA occurrence in Swiss forests was also contrary to the expectations we formed from Bugrova and Reznikova (1990) and Dekoninck et al. (2010). This could be a result of the intensities of recreational activities in Swiss forests all being below some critical thresholds the existence of which are suggested by the persistence of RWA in the disturbed plots of Vepsäläinen and Wuorenrinne (1978) and Bugrova and Reznikova (1990). Alternatively, such activities may be less disruptive to RWA than we originally considered. Similarly, the apparent lack of influence of proximities to roads and buildings was contrary to the expectations we formed regarding: proximity to roads from

Perron et al. (2019); and proximity to buildings from Wuorenrinne (1989) and Vepsäläinen and Wuorenrinne (1978). However, the results of Perron et al. (2019) concern the species *F. rufa* and *F. polyctena* which collectively comprised less than 7% of the observations of RWA mounds in our data. Given that all our data were collected within forests, it is also possible that any positive or negative effects of roads and buildings on RWA did not extend far enough into forests to be detectable at the locations from which our data were collected.

4.2 | Habitat associations of RWA

We found that RWA occurrence was positively associated with elevation and negatively associated with temperatures. This could be of concern in the context of predicted climate warming in Switzerland, particularly given the more extreme warming predicted for alpine regions (National Centre for Climate Services, 2018). If the lower temperatures of higher elevations are the main environmental characteristic that is favourable to RWA at such elevations, more extreme alpine warming would require range shifts by RWA to even higher elevations than would otherwise be necessary to find suitable habitats in a warmer climate. The potential requirement for such range shifts is of concern for two main reasons. Firstly, such range shifts would only be possible if the forest and invertebrate communities RWA rely on successfully make the same shifts. Secondly, there is uncertainty regarding both the rate at which alpine environments will warm and the rate at which RWA could colonise suitable habitat were it to become available. Consequently, there is a risk that RWA undergo local extinctions in Switzerland due to either RWA or the communities on which they rely failing to make the range shifts to higher elevations with sufficient speed to keep pace with climate warming. The NFI provides a good coverage of observations within the elevation range of 400 m to 2,000 m above sea level. Thus, our findings support the expectation that, in the parts of Europe outside Scandinavia, our two numerically dominant species (*F. lugubris* and *F. paralugubris*) are typically detected at locations with elevations greater than 1500m above sea level (Risch et al., 2016). While the final model from Vandegehuchte et al. (2017) for RWA occurrence did not include elevation or temperature as explanatory variables, it did include climate continentality which was assigned a positive coefficient. This was interpreted as an indirect effect of elevation because higher elevations in Switzerland are characterized by more continental climates. The variable selection procedure used by Vandegehuchte et al. (2017) included a LASSO regularization step. Supplied with a group of strongly correlated explanatory variables, LASSO regularization will typically assign only one of these variables a non-zero coefficient and the identity the variable thus selected from the group is not meaningful (Zou & Hastie, 2005). In contrast, we summarised each group of strongly correlated explanatory variables with a single synthetic variable. Considered in this context, we gain similar information about RWA associations with temperature and elevation from our models as we do from the selected model of

Vandegehuchte et al. (2017). In each of our studies, it seems that a group of correlated variables that included elevation, temperature and (where present) climate continentality were useful for predicting RWA occurrence. The LASSO step in the analysis of Vandegehuchte et al. (2017) appears to have resulted in climate continentality serving the purpose of this group in their selected model with other strongly correlated variables being assigned coefficients of zero. Our approach instead summarised the group containing elevation and temperature statistics with a single synthetic variable and this synthetic variable was found to be useful for predicting RWA occurrence. Both other broad spatial scale studies of RWA habitat associations did not consider the influence of elevation or temperature directly (Dekoninck et al., 2010; Procter et al., 2015). However, Sorvari (2016) identified climate warming as a key threat facing RWA.

Our finding of a positive association between RWA mound occurrence and conifers is supported by other broad spatial scale studies (Procter et al., 2015; Vandegehuchte et al., 2017) and the review of Risch et al. (2016) who report that RWA are typically found in forests either dominated by conifers or composed of a mixture of conifer and broad leaf trees. The association between RWA and conifers has multiple components: RWA incorporate conifer needles into their mounds when available (Domisch et al., 2008; Lenoir et al., 2001; Ohashi et al., 2007); conifers continue to provide sap for the phloem feeding mutualists of RWA when sap flow decreases in broadleaf trees after leaf-fall; and RWA benefit from the antifungal and antimicrobial properties of conifer resin which they incorporate into their mounds when available (Brütsch et al., 2017; Castella et al., 2008; Christe et al., 2003).

We detected a positive association between the amount of incoming solar radiation and RWA occurrence. This supports similar findings by Vandegehuchte et al. (2017) and those summarised in Risch et al. (2016). Risch et al. (2016) also emphasized the particular importance of insolation to smaller mounds, which may be more dependent upon direct warming by the sun in the cooler months during which ants are active. In addition, we found that increases in incoming solar radiation were generally associated with larger increases in the predicted probability of RWA occurrence at cooler, higher elevations than at warmer, lower elevations. RWA thermoregulate their mounds during their active periods (Jílková et al., 2015; Risch et al., 2016). Consequently, larger amounts of sunlight at higher elevations likely lead to lower energetic costs of thermoregulation for RWA mounds at such elevations. The positive association we detected between RWA occurrence and both gaps in the forest and openness in the forest canopy also supports the interpretation that warming by sunlight is important to RWA mounds. Gaps in the forest, described in terms of crown closure and canopy openness, were also found to be associated with RWA occurrence by Vandegehuchte et al. (2017) and *F. lugubris* occurrence by Punttila and Kilpeläinen (2009). The negative association between the proportion of open ground in units of habitat and *F. lugubris* occupancy detected by Procter et al. (2015) contrast with both these and our findings. However, this should not be extrapolated to a preference for continuous, completely closed canopy structures because the proportion of open ground had a

minimum value of 15% in the data collected by Procter et al. (2015). Similarly, our results should not be extrapolated to a preference for unvegetated, open ground because all our data came from habitats with sufficient vegetation to meet the NFI classifications of forest or shrub forest. Thus, there is agreement between our findings and those of Procter et al. (2015) in so far as RWA were detected in areas of vegetation with some openness even if the direction of the associations we detected with openness contrast. Furthermore, as reviewed by Risch et al. (2016), there is considerable variation among RWA species in apparent preferences for canopy closure. Canopy gaps also encourage the growth of ground cover vegetation, which can provide favourable habitat for RWA (Vandegehuchte et al., 2017). This association was strong enough in our data that gaps in the forest, canopy cover and cover by ground cover vegetation were grouped and summarised by a single synthetic variable during our dimension reduction step. This synthetic variable, NFI.S1.C13, was moderately correlated with elevation ($r = 0.56$), but only weakly correlated with any of our synthetic variables that described direct human disturbance of the forest ($|r| < 0.3$). This suggests that the forest characteristics described by NFI.S1.C13 were associated with elevation rather than forest management.

We found RWA occurrence to be negatively associated with temperatures during the driest 3 months of the year. In our data, these months typically intersected the period from late autumn to early spring during which RWA are dormant. Warmer temperatures during the dormant period disadvantage poikilothermic animals, such as RWA, by raising their metabolic rates and thereby forcing them to consume more of their winter energy stores (Sorvari, 2016). Thus, our results support concerns for the prospects of RWA in Switzerland in the context of the predicted warming of Swiss winters (National Centre for Climate Services, 2018).

Our results suggest an association between RWA mounds and soils with a lower potential wetness given rainfall as estimated by a soil wetness index. RWA excavate a system of chambers below their mounds which help to aerate the mound and serve as winter quarters (Frouz et al., 2016). Consequently, RWA would be disadvantaged by soils in which such chambers have the potential to be flooded. This could also explain the results of Punttila and Kilpeläinen (2009) who found *F. aquilonia* to be more frequent on mineral soils than in mires in Finland. In this study only *F. uralensis* was more frequently detected in mires than on mineral soils and this species was not present in our data. Thus, our results could be interpreted to be in accordance with those of Punttila and Kilpeläinen (2009) if interpreted in terms of indicators of the potential for below ground chambers to flood.

5 | CONCLUSIONS

We did not find evidence that RWA occurrence in Swiss forests was influenced by: forest management actions that produced disturbances less severe than complete habitat destruction; recreational use of the forest; proximity to roads; or proximity to buildings. However, we lacked observations from the same sites both before and after such human-related disturbances commenced. Nevertheless, the RWA data used in this study provide a baseline for comparison to the data that will be collected in future NFIs. Such comparisons will be important to investigate whether human-related disturbances are associated with RWA gains or losses over time across a broad spatial scale in Switzerland. This information will be highly valuable for RWA conservation and is currently absent from the literature. In the interim, we recommend the following strategies to aid the conservation of RWA in Switzerland: forests with high proportions of conifers should be conserved with priority given to those at higher elevations; and forests should be managed to remain open enough that sunlight can reach the level of ground cover vegetation with minimal disturbance to such vegetation. Our recommendations generally align with those of the other three broad spatial scale studies of RWA habitat associations. Dekoninck et al. (2010) recommended the conservation of existing RWA habitat and Vandegehuchte et al. (2017) recommended the conservation and establishment of coniferous forests. Procter et al. (2015) even found that plantations of non-native conifers can provide suitable additional habitat for *F. lugubris*. However, the establishment of additional native coniferous forests would be more conducive to the promotion of biodiversity. Both Vandegehuchte et al. (2017) and Dekoninck et al. (2010) recommended forest management that promotes varied vegetation structures and open canopies. Our recommendation regarding ground cover vegetation supports that of Vandegehuchte et al. (2017) but contrasts with that of Dekoninck et al. (2010). However, this contrasting recommendation is specific to forest edges in Belgium while our recommendations and those of Vandegehuchte et al. (2017) concern all forested areas of Switzerland. Finally, the negative associations we detected between RWA occurrence and temperatures raise concerns for the prospects of RWA in Switzerland in the context of the warmer winters predicted for all of Switzerland in coming decades and the more extreme climate warming predicted for alpine regions.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ddi.13217.

DATA AVAILABILITY STATEMENT

The code written to conduct the analysis is available from https://github.com/brfitzpatrick/rwaha. Swiss National Forest Inventory data may be requested directly from the Swiss National Forest Inventory. The digital elevation model (SwissALTI3D) and topographic landscape model (swissTLM3D) used in this work may be purchased from the Swiss Federal Office of Topography swisstopo. The Daymet climate interpolations for Switzerland are available upon request for non-commercial research after signing a usage contract.

ORCID

Benjamin R. Fitzpatrick https://orcid.org/0000-0003-1916-0939
Andri Baltensweiler https://orcid.org/0000-0003-1933-6535
Marielle Fraefel https://orcid.org/0000-0003-0992-6320
Anne Freitag https://orcid.org/0000-0002-9057-4475
Martijn L. Vandegehuchte https://orcid.org/0000-0003-1283-4654
Beat Wermelinger https://orcid.org/0000-0003-3235-6741
Anita C. Risch https://orcid.org/0000-0003-0531-8336

REFERENCES

Bivand, R. S., Pebesma, E., & Gómez-Rubio, V. (2013). *Applied spatial data analysis with R*, (2nd edn.). New York: Springer.

Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. http://dx.doi.org/10.1023/a:1010933404324

Brütsch, T., Jaffuel, G., Vallat, A., Turlings, T. C. J., & Chapuisat, M. (2017). Wood ants produce a potent antimicrobial agent by applying formic acid on tree-collected resin. *Ecology and Evolution*, 7(7), 2249–2254. https://doi.org/10.1002/ece3.2834

Bugrova, N. M., & Reznikova, J. I. (1990). The state of Formica polyctena FOERST. (Hymenoptera, Formicidae) population in recreation forests. *Memorabilia Zoologica*, 44, 13–19.

Busby, J. R. (1991). BIOCLIM - a bioclimate analysis and prediction system. *Plant Protection Quarterly*, 6(1), 8–9.

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A., Mace, G. M., Tilman, D., Wardle, D. A., Kinzig, A. P., Daily, G. C., Loreau, M., Grace, J. B., Larigauderie, A., Srivastava, D. S., & Naeem, S. (2012). Biodiversity loss and its impact on humanity. *Nature*, 486(7401), 59–67. https://doi.org/10.1038/nature11148

Castella, G., Chapuisat, M., Moret, Y., & Christe, P. (2008). The presence of conifer resin decreases the use of the immune system in wood ants. *Ecological Entomology*, 33(3), 408–412. https://doi.org/10.1111/j.1365-2311.2007.00983.x

Chavent, M., Kuentz-Simonet, V., Liquet, B., & Saracco, J. (2012). ClustOfVar: An R Package for the Clustering of Variables. *Journal of Statistical Software*, 50(13), 1–16. http://dx.doi.org/10.18637/jss.v050.i13

Christe, P., Oppliger, A., Bancalà, F., Castella, G., & Chapuisat, M. (2003). Evidence for collective medication in ants. *Ecology Letters*, 6(1), 19–22. https://doi.org/10.1046/j.1461-0248.2003.00395.x

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., & Böhner, J. (2015). System for automated geoscientific analyses (SAGA) v. 2.1.4. *Geoscientific Model Development*, 8, 1991–2007. https://doi.org/10.5194/gmd-8-1991-2015

Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, 88(11), 2783–2792. https://doi.org/10.1890/07-0539.1

Dekoninck, W., Hendrickx, F., Grootaert, P., & Maelfait, J. (2010). Present conservation status of red wood ants in north-western Belgium: Worse than previously, but not a lost cause. *European Journal of Entomology*, 107(2), 209–218. https://doi.org/10.14411/eje.2010.028

Domisch, T., Ohashi, M., Finér, L., Risch, A. C., Sundström, L., Kilpeläinen, J., & Niemelä, P. (2008). Decomposition of organic matter and nutrient mineralisation in wood ant (Formica rufa group) mounds in boreal coniferous forests of different age. *Biology and Fertility of Soils*, 44, 539–545. https://doi.org/10.1007/s00374-007-0248-0

Domisch, T., Risch, A. C., & Robinson, E. J. H. (2016). Wood ant foraging and mutualism with aphids. In J. A. Stockan, & E. J. H. Robinson (Eds.), *Wood ant ecology and conservation* (pp. 145–176). Cambridge: Cambridge University Press.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1), 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Düggelin, C., & Keller, M. (2017). *Schweizerisches Landesforstinventar Feldaufnahme-Anleitung*. Retrieved from www.lfi.ch

ESRI (2010a). *ArcGIS desktop*. Environmental Systems Research Institute.

ESRI (2010b). *ArcMap desktop*. Environmental Systems Research Institute.

Evans, I. S. (1972). General geomorphometry, derivatives of altitude, and descriptive statistics. In R. J. Chorley (Ed.), *Spatial analysis in geomorphology* (pp. 17–90). New York: Harper & Row.

Federal Forest Act (1991). *SR 921.0*. Retrieved from https://www.admin.ch/opc/en/classified-compilation/19910255/index.html

Freitag, A., Dischinger, C., & Cherix, D. (2008). Formica pratensis (Hyménoptères: Formicidae) dans le canton de Vaud: état des peuplements et importance des talus de routes comme milieu de substitution. *Bulletin de la Société Vaudoise des Sciences Naturelles*, 91(1), 47–68. http://doi.org/10.5169/seals-282144

Frouz, J., Braunisch, V., & Sorvari, J. (2016). Contribution of wood ants to nutrient cycling and ecosystem function. In J. A. Stockan, & E. J. H. Robinson (Eds.), *Wood ant ecology and conservation* (pp. 207–220). Cambridge: Cambridge University Press.

Fu, P., & Rich, P. M. (2002). A geometric solar radiation model with applications in agriculture and forestry. *Computers and Electronics in Agriculture*, 37, 25–35. https://doi.org/10.1016/S0168-1699(02)00115-1

Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2015). Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *Journal of Computational and Graphical Statistics*, 24(1), 44–65. https://doi.org/10.1080/10618600.2014.907095

Hothorn, T., Bühlmann, P., Dudoit, S., Molinaro, A., & Van Der Laan, M. J. (2006). Survival ensembles. *Biostatistics*, 7(3), 355–373. https://doi.org/10.1093/biostatistics/kxj011

Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*, 15(3), 651–674. https://doi.org/10.1198/106186006X133933

IUCN (2019). *The IUCN red list of threatened species*. Retrieved from http://www.iucnredlist.org

Iverson, L. R., Dale, M. E., Scott, C. T., & Prasad, A. (1997). A GIS-derived integrated moisture index to predict forest composition and productivity of Ohio forests (U.S.A.). *Landscape Ecology*, 12(5), 331–348. https://doi.org/10.1023/A:1007989813501

Janitza, S., Strobl, C., & Boulesteix, A. L. (2013). An AUC-based permutation variable importance measure for random forests. *BMC Bioinformatics*, 14(1), 119. https://doi.org/10.1186/1471-2105-14-119

Jílková, V., Cajthaml, T., & Frouz, J. (2015). Respiration in wood ant (Formica aquilonia) nests as affected by altitudinal and seasonal changes in temperature. *Soil Biology and Biochemistry*, *86*, 50–57. https://doi.org/10.1016/j.soilbio.2015.03.024

Jurgensen, M. F., Finér, L., Domisch, T., Kilpeläinen, J., Punttila, P., Ohashi, M., Niemelä, P., Sundström, L., Neuvonen, S., & Risch, A. C. (2008). Organic mound-building ants: Their impact on soil properties in temperate and boreal forests. *Journal of Applied Entomology*, *132*(4), 266–275. https://doi.org/10.1111/j.1439-0418.2008.01280.x

Keller, M. (Ed.) (2013). *Umfrage Anleitung LFI4*. Retrieved from www.lfi.ch

Laine, K. J., & Niemela, P. (1980). The influence of ants on the survival of mountain birches during an Oporinia autumnata (Lep., Geometridae) outbreak. *Oecologia*, *47*, 39–42. https://doi.org/10.1007/BF005 41773

Lenoir, L., Persson, T., & Bengtsson, J. (2001). Wood ant nests as potential hot spots for carbon and nitrogen mineralisation. *Biology and Fertility of Soils*, *34*(4), 235–240. https://doi.org/10.1007/s003740100405

Maňák, V., Nordenhem, H., Björklund, N., Lenoir, L., & Nordlander, G. (2013). Ants protect conifer seedlings from feeding damage by the pine weevil Hylobius abietis. *Agricultural and Forest Entomology*, *15*(1), 98–105. https://doi.org/10.1111/j.1461-9563.2012.00597.x

McCune, B., & Keon, D. (2002). Equations for potential annual direct incident radiation and heat load. *Journal of Vegetation Science*, *13*(4), 603–606. https://doi.org/10.1111/j.1654-1103.2002.tb02087.x

National Centre for Climate Services (2018). *CH2018 – Climate Scenarios for Switzerland*, *Technical Report*, Zurich: (A. M. Fischer & K. M. Strassmann, Eds.). National Centre for Climate Services. https://www.nccs.admin.ch/nccs/en/home/climate-change-and-impacts/swiss-climate-change-scenarios/technical-report.html

Nix, H. A. (1986). A biogeographic analysis of Australian elapid snakes. In R. Longmore (Ed.), *Atlas of Elapid Snakes of Australia* (pp. 4–15). Canberra: Australian Government Public Service.

Ohashi, M., Kilpeläinen, J., Finér, L., Risch, A. C., Domisch, T., Neuvonen, S., & Niemelä, P. (2007). The effect of red wood ant (Formica rufa group) mounds on root biomass, density, and nutrient concentrations in boreal managed forests. *Journal of Forest Research*, *12*(2), 113–119. https://doi.org/10.1007/s10310-006-0258-z

Perron, M., Glanzmann, I., & Freitag, A. (2019). Habitatselektion von zwei Waldameisenarten (Formica rufa und F. polyctena). *Swiss Forestry Journal*, *170*(1), 32–39. https://doi.org/10.3188/szf.2019.0032

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., Raven, P. H., Roberts, C. M., & Sexton, J. O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, *344*(6187), 1246752. https://doi.org/10.1126/science.1246752

Procter, D. S., Cottrell, J., Watts, K., & Robinson, E. J. H. (2015). Do non-native conifer plantations provide benefits for a native forest specialist, the wood ant Formica lugubris? *Forest Ecology and Management*, *357*, 22–32. https://doi.org/10.1016/j.foreco.2015.07.034

Punttila, P., & Kilpeläinen, J. (2009). Distribution of mound-building ant species (Formica spp., Hymenoptera) in Finland: Preliminary results of a national survey. *Annales Zoologici Fennici*, *46*(1), 1–15. https://doi.org/10.5735/086.046.0101

R Core Team (2018). *R: A Language and Environment for Statistical Computing*, Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Risch, A. C., Ellis, S., & Wiswell, H. (2016). Where and why? Wood ant population ecology. In J. A. Stockan, & E. J. H. Robinson (Eds.), *Wood Ant Ecology and Conservation* (pp. 81–105). Cambridge: Cambridge University Press.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*(8), 913–929. https://doi.org/10.1111/ecog.02881

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*(77), 1–8. https://doi.org/10.1186/1471-2105-12-77

Robinson, E. J. H., Stockan, J. A., & Iason, G. R. (2016). Wood ants and their interaction with other organisms. In J. A. Stockan, & E. J. H. Robinson (Eds.), *Wood ant ecology and conservation* (pp. 177–206). Cambridge: Cambridge University Press.

Sorvari, J. (2016). Threats, conservation and management. In J. A. Stockan, & E. J. H. Robinson (Eds.), *Wood ant ecology and conservation* (pp. 264–286). Cambridge: Cambridge University Press.

Stockan, J. A., Robinson, E. J. H., Trager, J. C., Yao, I., & Seifert, B. (2016). Introducing wood ants: Evolution, phylogeny, identification and distribution. In J. A. Stockan, & E. J. H. Robinson (Eds.), *Wood ant ecology and conservation* (pp. 1–36). Cambridge: Cambridge University Press.

Strobl, C., Boulesteix, A. L., Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics*, *9*(1), 307. https://doi.org/10.1186/1471-2105-9-307

Strobl, C., Boulesteix, A. L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, *8*(1), 25. https://doi.org/10.1186/1471-2105-8-25

Thornton, P. E., Running, S. W., & White, M. A. (1997). Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology*, *190*(3–4), 214–251. https://doi.org/10.1016/S0022-1694(96)03128-9

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, *58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: A retrospective. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, *73*(3), 273–282. https://doi.org/10.1111/j.1467-9868.2011.00771.x

Tilman, D., Isbell, F., & Cowles, J. M. (2014). Biodiversity and ecosystem functioning. *Annual Review of Ecology, Evolution, and Systematics*, *45*(1), 471–493. https://doi.org/10.1146/annurev-ecolsys-12021 3-091917

Valiente-Banuet, A., Aizen, M. A., Alcántara, J. M., Arroyo, J., Cocucci, A., Galetti, M., García, M. B., García, D., Gómez, J. M., Jordano, P., Medel, R., Navarro, L., Obeso, J. R., Oviedo, R., Ramírez, N., Rey, P. J., Traveset, A., Verdú, M., & Zamora, R. (2015). Beyond species loss: The extinction of ecological interactions in a changing world. *Functional Ecology*, *29*(3), 299–307. https://doi.org/10.1111/1365-2435.12356

Vandegehuchte, M. L., Wermelinger, B., Fraefel, M., Baltensweiler, A., Düggelin, C., Brändli, U.-B., Freitag, A., Bernasconi, C., Cherix, D., & Risch, A. C. (2017). Distribution and habitat requirements of red wood ants in Switzerland: Implications for conservation. *Biological Conservation*, *212*, 366–375. https://doi.org/10.1016/j.biocon.2017.06.008

Vepsäläinen, K., & Wuorenrinne, H. (1978). Ecological effects of urbanization on the mound-building Formica L. species. *Memorabilia Zoologica*, *29*, 191–202.

Weiss, A. D. (2001). In Topographic position and landforms analysis. *ESRI user conference*, San Diego: ESRI.

Wuorenrinne, H. (1989). Effects of urban pressure on colonies of Formica rufa group (Hymenoptera, Formicidae) in the town of Espoo (Finland). *Annales Zoologici*, *42*(15), 335–344.

Xu, T., & Hutchinson, M. (2011). *ANUCLIM*, Canberra: The Australian National University.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, *67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

BIOSKETCH

Benjamin R. Fitzpatrick works at the interface between ecology, statistics and machine learning. Ben's work on fishes, soil carbon and ants has all involved the use of models to understand and predict aspects of the natural world. In recent years, Ben has focused on predicting spatial distributions and understanding the processes influencing these distributions. Ben's interest in cautiously harnessing the predictive power of machine learning methods without sacrificing model interpretability is complemented by his interests in model and data visualization. More information about Ben's work and interests can be found at his website: https://brfitzpatrick.github.io/.

Author Contributions: ACR, BW, MLV and BRF conceived and designed the study. BRF designed and conducted the statistical analysis; performed all statistical programming; and produced all graphics. BRF wrote the original manuscript and all authors contributed to its revision. BRF and CD selected and prepared the NFI derived data for analysis. AB calculated the DEM derived data. BRF and AB selected and prepared the DEM derived data for analysis. MF calculated the distances to roads and buildings. BRF and MF prepared the distances to roads and buildings for analysis. BRF calculated and prepared the Daymet derived data for analysis. AF supervised the identification of the ant specimens. BRF prepared the ant occurrence data for analysis.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.