

¹ Plague risk in the western United States over seven decades of
² environmental change

³ Colin J. Carlson^{1,†}, Sarah N. Bevins², and Boris V. Schmid³

⁴ ¹*Center for Global Health Science and Security, Georgetown University Medical Center,
5 Washington, D.C. 20057, USA*

⁶ ²*US Department of Agriculture Animal and Plant Health Inspection Service–Wildlife Services
7 National Wildlife Research Center, Fort Collins, Colorado, USA*

⁸ ³*Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo,
9 NO-0316 Oslo, Norway*

¹⁰ [†]*Correspondence should be directed to colin.carlson@georgetown.edu*

¹¹ Last updated on February 27, 2021

¹²

Abstract

¹³

After several pandemics over the last two millennia, the wildlife reservoirs of plague (*Yersinia pestis*) now persist around the world, including in the western United States. Routine surveillance in this region has generated comprehensive records of human cases and animal seroprevalence, creating a unique opportunity to test how plague reservoirs are responding to environmental change. Here, we develop a new method to detect the signal of climate change in infectious disease distributions, and test whether plague reservoirs and spillover risk have shifted since 1950. We find that plague foci are associated with high-elevation rodent communities, and soil biochemistry may play a key role in the geography of long-term persistence. In addition, we find that human cases are concentrated only in a small subset of endemic areas, and that spillover events are driven by higher rodent species richness (the amplification hypothesis) and climatic anomalies (the trophic cascade hypothesis). Using our detection model, we find that due to the changing climate, rodent communities at high elevations have become more conducive to the establishment of plague reservoirs—with suitability increasing up to 40% in some places—and that spillover risk to humans at mid-elevations has increased as well, although more gradually. These results highlight opportunities for deeper investigation of plague ecology, the value of integrative surveillance for infectious disease geography, and the need for further research into ongoing climate change impacts.

31 Introduction

32 The distribution and burden of infectious diseases will be entirely reshaped by global environmental change. Scientific consensus suggests that over the next century, the combined effect of climate change, land degradation and transformation, and increasing human-wildlife contact will bring about a massive increase in the spillover of pathogens that originate in wildlife (zoonotic diseases)^{1,2} and the burden of infections transmitted by arthropods (vector-borne diseases)^{3,4}. While there is substantial research efforts working to project these future changes, the impacts of current environmental change on infectious disease burden in the world today is underexplored. Based on current evidence, land use change is the best-supported leading driver of zoonotic emergence^{5,6}; much less is known about climate change impacts to date. This is due, in large part, to methodological limitations: the “detection and attribution” methods that are best suited to this problem require substantial data on disease prevalence or incidence over extensive periods, as well as complicated model designs (e.g., counterfactual climate scenarios without climate change)^{7,8}.

45 Instead, many projections of climate change impacts rely on *ecological niche models* (also known as species distribution models), a set of regression and machine learning approaches that relate climate to the geographic range of a species^{9,10}. Usually, these approaches are an oversimplification of reality, especially for pathogens: for example, a map of anthrax (*Bacillus anthracis*) may classify west Texas as an endemic zone, even though the system is characterized by epizootics that are sometimes years apart¹¹. Ecological niche models are therefore an imperfect tool for exploring climate change impacts. These methods work well for mapping current distributions, for projecting single-time-slice distributions under future climates, and – in some recent work – for projecting continuous-time change¹². Retrospective work to reconstruct climate change impacts is much rarer, and is usually restricted to work that builds two species distribution models for contrasting time intervals and compares them^{13,14,15}. This approach is less than ideal, forcing researchers to violate the assumption that species’ geographic ranges are at equilibrium^{16,17}; to aggregate data into somewhat arbitrary time periods; and to compare models trained on non-independent but non-overlapping datasets, which will generate different biological response curves simply because of model uncertainty. In this framework, it is also difficult to eliminate alternate hypotheses for why a species’ apparent distribution might change, like noise in the detection process or shifting abundance patterns.

63 Recently, a growing set of tools have tried to grapple with the temporal variability exhibited by the distribution of infectious diseases. Though most disease maps are treated as the long-term average of temporally-dynamic processes, *time-specific ecological niche modeling* has been proposed as an alternative that captures the dynamic nature of transmission. 66 Almost always, though, these methods have been implemented at the finest temporal scales: monthly¹⁸ or seasonal^{19,20}. As of yet, this approach has been mostly untested as a way 68 of understanding disease distributions over multiple years—and ideally, of contextualizing 70 the impacts of environmental change over decades (but see¹⁵).

71 Few systems provide a better opportunity to test this approach than plague, a globally-
72 cosmopolitan zoonotic infection caused by the bacterium *Yersinia pestis*. The global distri-
73 bution of plague has been far from stable over the past two centuries; the Third Pandemic
74 (late 18th to mid-20th Century)^{21,22,23,24,25} in particular was responsible for the introduc-
75 tion of *Y. pestis* into many new regions that were environmentally suitable but otherwise
76 uncolonized, particularly the Americas^{26,27,28}. In some of these regions, outbreaks have
77 faded over time, while in others, plague foci have persisted and the pathogen has become
78 endemic, maintained by a sylvatic cycle in rodent reservoirs and flea vectors²⁸. Rodent
79 biodiversity hotspots may be particularly conducive to the formation of these reservoirs²⁹,
80 a possible case of *biodiversity amplification* effects^{30,31}, where the diversity of competent
81 hosts allows a virulent pathogen to be maintained at more stable levels. Though under-
82 explored, emerging evidence also suggests that plague may persist in the soil, possibly by
83 acting as an endosymbiont with amoebas³², from which it sporadically can reinfect burrow-
84 ing rodents³³. Soil conditions may therefore further constrain the distribution of plague
85 reservoirs^{34,35,36,37}. Like other pathogens that can persist in the soil^{38,39}, provisional evi-
86 dence suggests that plague may be limited by soil salinity^{40,41}, soil organic carbon, and
87 alkalinity⁴². Though these factors may have limited influence in the short-term dynamics
88 of plague in any one location, at continental scales, they could reasonably be expected to
89 shape where plague foci have become established.

90 Both experimental and ecological analysis suggests that plague dynamics are also highly
91 sensitive to climatic conditions. The disease's sensitivity to bioclimatic conditions has been
92 documented throughout its life cycle, but is particularly pronounced on the arthropod level,
93 where temperature (and to a lesser degree humidity) influence the rate at which various
94 flea species move through their life cycle⁴³. Flea species differ in their temperature sensi-
95 tivities⁴³, making the local composition of flea communities an important consideration,
96 as well as in their ability to transmit plague, either through early-phase transmission or
97 blockage-induced transmission^{44,43,45}. The bacterium appears able to rapidly evolve its
98 ability to favor one over the other transmission mode⁴⁶ (or maintains a standing varia-
99 tion in its extended phenotype within populations). Temperature also directly influences
100 biochemical aspects of the transmission efficiency of the plague bacterium, particularly
101 when temperatures rise above 27°C⁴⁴, presumably by negatively influencing the stability
102 of the biofilm that the bacterium forms in fleas. Temperature also finally influences rodent
103 populations, including through a mechanism generally referred to as a *trophic cascade*: cli-
104 matic anomalies influence primary productivity, driving changes in rodent density, which
105 in turn change the density and biting preferences of fleas^{47,48}. The combination of these
106 environmental sensitivities, when playing out across the scale of ecosystems, can lead to
107 widespread synchronicities in plague epizootic periods⁴⁹.

108 All of these lines of evidence suggest that plague should be broadly sensitive to en-
109 vironmental change, and that in systems where trends in plague occurrence have been
110 tracked, an anthropogenic signal might be detectable. The United States is the perfect
111 system to test this approach, as data in this region are particularly abundant; human case

112 data goes back over a century, to plague's first introduction on the Pacific coast in early
113 20th Century^{50,51,26}. Moreover, the U.S. Department of Agriculture has collected plague
114 seropositivity data from wildlife for multiple decades through the USDA National Wildlife
115 Disease Program⁵². Combined, these national datasets include more records than many
116 global studies of pathogen distributions³⁸, making this system an ideal testing ground. To-
117 gether, these data also cover nearly a century of environmental change, a temporal scope
118 that allows time-specific ecological niche modeling to be implemented. This also allows us
119 to revisit one of the only previous attempts at this approach, which compared models of
120 plague risk in the western U.S. based on case data in three multiyear time slices (1965-69,
121 1980-84, and 1995-99), and concluded that plague risk had expanded since 1950 and would
122 continue to do so in the future¹⁵.

123 In this study, we revisit this prediction by using two independent data streams (human
124 cases and wildlife serology) in a machine learning model called *Bayesian additive regression*
125 *trees* (BART)⁵³ (see Methods for a detailed explanation). Climatic reconstructions are
126 readily available for the duration of our study (1950–2017), allowing us to use annual
127 climate layers (including long-term anomalies) corresponding to the year of each plague
128 case. This pairing allows us to improve model precision relative to long-term averages,
129 to differentiate areas of ephemeral versus persistent risk, and to identify the fingerprint
130 of environmental change in risk trends. We also test whether the distribution of plague
131 in this region is responsive to rodent biodiversity or soil chemistry and macronutrients,
132 offering detailed insights into the factors that maintain plague risk. Finally, we propose
133 a new approach that harnesses *BART with random intercepts* (riBART) to account for
134 historical variability in detection and sampling, allowing us to confidently identify the
135 signal of changing environmental conditions in plague prevalence over time. In doing so,
136 we propose the first extension of ecological niche modeling that nods towards the ultimate
137 aim of detection and attribution of anthropogenic climate change impacts on the geographic
138 distribution of infectious diseases.

139 Results

140 The distribution of plague

141 We generated two primary models of plague over time. The first covered 9,761 animals
142 sampled for plague (2000 to 2017), and performed well (training AUC = 0.836; Extended
143 Data Figure 1). The second covered a total of 430 human cases of plague (1950 to 2005), and
144 performed very well (training AUC = 0.909; Extended Data Figure 2). When both models
145 were rerun with an overlapping “test period” of 2000 to 2005 withheld, they performed
146 adequately, with the human model (AUC = 0.820) performing better than the wildlife
147 model (AUC = 0.775). As both models performed well in temporal cross-validation, we
148 used both to make annual predictions from 1950 to 2017, and split predictions into binary
149 presence or absence risk maps for each year using the true skill statistic.

150 Both models found that the majority of plague risk in the western United States is, as
151 expected, found west of the 100th meridian (Figure 1). The human model mostly predicts
152 risk in the “Four Corners” region (Utah, Colorado, Arizona, and New Mexico), where that
153 risk is relatively stable across years. In contrast, the wildlife model predicts risk fairly
154 expansively west of 100°, including at much higher latitudes than the human risk model.
155 Risk varies much more across years in this model, but several areas are predicted to be
156 environmentally suitable across years from Montana to west Texas. The suitable areas
157 identified by the wildlife model in the southwest are less uniform than the human model,
158 likely reflecting a finer-scale differentiation of risk. There are two main reasons the human
159 model might discriminate less in this region: human cases may be reported in different
160 locations than the site of initial spillover, and occurrence points were randomly resampled
161 at the county level (as data have been previously de-identified).

162 For the most part, we found that risk areas identified in the human model were a
163 subset of the much broader predictions made by the wildlife model (Figure 2), with three
164 major exceptions. The human model identified much broader risk in southern Arizona and
165 New Mexico, likely due to how the cases were randomized at county levels. The human
166 model also predicted areas of risk further east, in regions like east Texas or Oklahoma
167 where plague is not known to be endemic (and, in this regard, the wildlife model better
168 captures the known distribution), but conditions may be broadly favorable. Finally, and
169 most notably, the human model predicted plague risk throughout California, in places
170 that have previously been identified as high-risk⁵⁴. This likely reflects a deficit of data
171 from Californian sources in our wildlife model, as state wildlife surveillance is curated
172 independently. Together, our findings indicate the value of comprehensive surveillance,
173 and the possibility that zoonotic reservoirs may be more expansive than areas of known
174 spillover.

175 Ecological insights

176 Our models identified a number of intersecting factors that maintain plague reservoirs and
177 create the right conditions for spillover, many of which have been previously identified
178 by ecological and epidemiological models (Extended Data Figures 3, 4, 5, 6, 7, 8, 9). A
179 handful of factors are important in both animal and human spillover models, and have
180 similar response profiles: elevation, with higher risk at higher elevations; rodent species
181 richness, with a similar positive effect; and the sodium (Na) and calcium (Ca) content of
182 the soil, both with a negative effect on plague risk. As these factors are shared between
183 the models, we can tentatively conclude that these factors relate to what happens in the
184 wildlife, and are not substantially altered by the additional spillover process that the hu-
185 man model incorporates. A fourth environmental factor that is significant in both models
186 is temperature, but with different response profiles in how means, maxima and anomalies
187 in temperature affected the risk of plague. Finally, we found strong effects of clay and iron
188 content of the soil, which are shared between the two models but follow different profiles,

189 as well as the sandiness of the soil (wildlife model only) and the variability in annual pre-
190 cipitation within the area (human model only). A list of all variables and abbreviations is
191 given in Extended Data Table 1.

192

193 **Elevation:** Both models indicate that plague risk increases at higher elevations, particu-
194 larly above 2,000 meters, compatible with previous findings in this system^{54,55,56,57}. Using
195 spatial partial dependence plots, we were able to show that the abrupt transition in plague
196 suitability at the 100th meridian (100° W) was driven by the elevational layer in both mod-
197 els, and not suitably explained by any bioclimatic factors (Extended Data Figures 10, 11).
198 Elevation has also been previously associated with plague on other continents^{42,58,59,60,61,62},
199 and while the general trend is that there is a lower threshold elevation (and an upper limit,
200 at the extreme altitudes in the Himalayas), that threshold differs substantially between
201 countries. For example, Brazil's plague reservoirs start at 500 meters above sea level, and
202 are limited by the landscape to no more than 1,000 meters⁶³, while Madagascar's urban
203 plague reservoir in Mahajanga is at sea level⁶⁴, as are the plague reservoirs surrounding
204 the Caspian Sea in Central Asia⁶⁵. Elevation therefore seems to represent the local ecology
205 and distributional limits of fleas and rodents, rather than a global proxy for a bioclimatic
206 or atmospheric variable (e.g. partial CO₂ pressure).

207

208 **Rodent richness:** Likewise, both models found a higher suitability for plague in areas
209 with higher rodent richness, with the factor being only second to elevation in importance
210 in the human plague-risk model. This points to the possibility that high-elevation hotspots
211 of rodent biodiversity may help maintain enzootic plague transmission, a possible case of
212 the biodiversity amplification effect that has also been found by similar work in China²⁹.
213 As in China, it is unclear whether the increased biodiversity itself has a positive effect
214 on plague maintenance, or whether it merely signals an increased chance that certain key
215 rodent species (particularly synanthropic ones) are locally present. If there are positive,
216 general associations between rodent diversity and plague risk, this would be an exception
217 to widespread evidence of biodiversity dilution effects for other vector-borne diseases⁶⁶.
218 Most theoretical models of the dilution effect rely on a skewed distribution of host com-
219 petence, where higher host diversity leads to reduced transmission in the narrow subset of
220 competent hosts^{67,68,69}; plague is perhaps uniquely capable of infecting and causing dis-
221 ease in hundreds of host species⁷⁰, and therefore may not exhibit these dynamics, though
222 only a limited number of species develop a high enough viremia to infect a feeding flea.
223 Alternately, it may be that scale underlies this pattern; theory suggests that dilution ef-
224 fects are strongest at small scales, while amplification effects may be normal at continental
225 scales^{30,71}. Finally, it might simply be that plague behaves differently than other vector-
226 borne diseases because plague can also spread through pneumonic transmission and prey
227 consumption, which produce different dynamics.

228

229 **Climate:** Temperature was a universally-important predictor across all models, while

surprisingly, precipitation only minimally influenced predictions. In the wildlife model, we found a negative relationship between mean temperature and plague risk—an unusual response curve for a vector-borne disease⁷². Plague across the globe occurs in a wide variety of ecosystems, from the tropics in Africa, to semi-arid deserts in Kazakhstan and the high mountains in Central Asia, and appears able to persist across a large temperature range. Its apparent association with colder locations in the United States may therefore not be directly related to some temperature threshold, but possibly to the ecological niche of key maintenance hosts or key vector species. For human plague risk, we see a sharp increase as the mean and maximum annual temperature falls between 5° and 14°, respectively, matching previous findings from empirical work on North American fleas^{73,74} and global models of the Third Pandemic spread⁷⁵. This may reflect the underlying thermal ecology of the pathogen⁷², or potentially some combination of rodent habitat and human behavior (e.g., these thresholds might delineate the general type of wilderness areas in the Rocky Mountains where people live alongside plague reservoirs).

In addition to the effect of long-term climatic averages, the temporal structure of the model allowed us to detect a strong effect of interannual variability. In the wildlife model, we observed an increase in plague prevalence during anomalously warm years, a result that has been previously reported for semi-arid desert ecosystems⁷⁶, as well as for human cases in the United States⁵¹. Warmer years are likely to increase rodent density, both directly through mild winters⁷⁷ and indirectly through higher primary productivity; flea populations in turn tend to follow rodent density, with some degree of lag^{77,78,79}. In contrast, in the human model, spillover was most likely in anomalously wet, cold years. This matches previous findings in other systems^{78,79,80}, which have been attributed to another kind of trophic cascade: when seasonal fluctuations become unfavorable to rodent populations after a recent high, and these rodent populations contract, fleas aggregate on the remaining rodents, both facilitating the dissemination of plague between rodents, and making fleas more eager to seek secondary hosts to feed on, thus leading to increased spillover risk.

Soil: Finally, we found that both models provided evidence that the long-term persistence of plague foci is related to properties of the soil. Our modeling suggests that *in vivo*, soils with higher proportions of sand and intermediate proportions of clay (~20-30%) (Extended Data Figure 8), low sodium and calcium contents, and mid-to-low concentrations of iron seem to be most conducive to plague. Although not included in either model after variable set reduction, we also found that soil pH may limit persistence, with more alkaline soils favored in the wildlife model. The observed response curves are somewhat unusual, given that both human and wildlife cases peaked in the raw data around a soil pH of 8.2-8.4 (Extended Data Figure 12); it may be that this reflects colinearities with other soil traits, or simply a smaller effect of pH compared to other soil characteristics.

The role of a soil compartment in the maintenance of plague reservoirs has been under consideration for more than a century, and various aspects of a soil-cycle of plague have been independently confirmed in laboratory settings. These include survival in the soil for

months to years in a laboratory setting, either in association with amoebas (*Acanthamoeba castellanii* and *Dictyostelium discoideum*) or independently^{34,32,36,81}; the existence of *Y. pestis* in soil in wildlife plague foci⁴⁰; the sporadic return from soil into a rodent population³³; and geographic correlations between plague foci and various soil properties^{42,59,40}. Mechanisms through which these factors affect plague foci may be directly related to the bacterium, or through soil factors that influence the vector or the host. Fleas living in burrows, for example, are negatively affected in all aspects of their lifecycle (fecundity, development, survival, and activity) in environments with a 100-fold higher level of fractional CO₂ than the atmospheric fraction⁸². This link to ventilation can explain the role of soil characteristics: a poorly ventilated burrow (impermeable, non-sandy soil) may reach up to 65-fold higher level of fractional CO₂, whereas a unventilated permeable (sandy) soil would only reach up to 25-fold higher levels, and a well-ventilated permeable soil would only reach up to 7-fold higher⁸³. Likewise, soil mineral content could have downstream effects on the homeostasis of virulence-related minerals in the body of the host. As a parallel, one study in Brazil found that cattle grazing on iron-rich soil would lead to highly lethal infections of normally non-pathogenic *Yersinia pseudotuberculosis*, presumably by perturbing the ability of the host to bind iron away from being utilized by the bacterium during an infection⁸⁴. Both iron^{85,86,87} and calcium^{88,89} are virulence factors of *Y. pestis*, and we tentatively hypothesize that plague foci persist better in regions where *Y. pestis* is not facilitated by the environment to be highly virulent.

Detecting environmental impacts on change over time

We found a strong temporal trend in both climate and plague risk since 1950 (Figure 3). To test whether this signal might be confounded by exogenous factors, we used a novel approach where we trained BART models with random intercepts (riBART) for each year, and projected the model again without random effects over the 68 year period (see Methods). The random intercepts identified interannual variation in prevalence (Extended Data Figures 13,14), detrended the data, and allowed the models to identify climate signal minus the confounder without substantially changing overall predictions (Extended Data Figure 15). Subsequently, we predicted how suitability changed using the same model without random intercepts; this allowed us to be confident that the changing suitability we identified in these “detection models” was the consequence of constant relationships between temperature, precipitation, and plague transmission.

Both detection models identified a meaningful signal of temporal variation. The random intercepts identified a signal of rising prevalence through the wildlife dataset, particularly increasing after 2011 when the diagnostics were changed (Extended Data Figures 13,14). In the human model, we identified a much more subtle long-term quadratic trend peaking in the 1980s, matching a pattern that has been previously attributed to climate cycles like the Pacific Decadal Oscillation.⁵¹ Surprisingly, the “detection” models identified an even stronger pattern of change over time (Extended Data Figure 16). In the wildlife model,

310 suitability increased an average of 4.8%, and 4.9% in the detection model, with a much
311 fatter tail to the distribution as well. In the human model, suitability increased an average
312 of 1.7% from 1950 to 2017, and 2.1% in the detection model. In much of the region, we
313 found that plague suitability increased by 30 to 40% over the entire interval. We found that
314 suitability rose most substantially in the wildlife model at high elevations, while spillover
315 risk increased more gently, and peaked at mid-elevations (Figure 4, Extended Data Figure
316 17). Because the detection models only predict change across years based on temperature
317 and precipitation, we are confident that the increase in the long term signal of warming
318 (roughly 0.8° in the region since 1950) and higher anomalous precipitation are responsible
319 for these predicted changes (see Extended Data Figures 18,19). We conclude that, even
320 with several confounding factors, environmental change since the 1950s may have helped
321 plague reservoirs become established at higher elevations, and slightly increased the risk of
322 spillover into human populations at mid-elevations. Over the coming half-century, previous
323 work suggests that the shift towards higher elevations is likely to continue in this region⁵⁴.

324 Discussion

325 Our study shows that human and wildlife data can be used together to map plague reser-
326 voirs and spillover risk in the United States, and to make meaningful inferences about
327 ecological drivers of transmission. We found support for two major hypotheses: the biodi-
328 versity amplification effect and the trophic cascade hypothesis. Support for these patterns
329 has increasingly been found across systems, and points to a view of plague risk where
330 weather conditions (and their impact on flea vectors) in rodent biodiversity hotspots are
331 the primary driver of transmission and spillover. We further found strong evidence that the
332 North American distribution of plague is heavily influenced by soil conditions. The global
333 distributions of soil-persistent bacteria like anthrax (*Bacillus anthracis*), tularemia (*Francisella tularensis*), and botulism (*Clostridium botulinum*) are known to be constrained by
334 the biochemical properties of soil. Less is known about plague, which is not spore-forming,
335 and until recently was mostly thought to behave like a typical vector-borne zoonosis. It
336 may be that soil properties affect the suitability of burrows for higher flea densities, or
337 determine host homeostasis for minerals that impact the virulence of the infection; plague
338 foci might therefore fall in the narrow range of conditions that can harbor higher densities
339 of fleas, but do not substantially increase the lethality of the infection. However, increasing
340 evidence also suggests that the bacterium can persist in the soil, possibly through symbi-
341 otic relationships with amoebas, for weeks to months—and possibly even years^{37,35}. These
342 complexities underscore the importance of a One Health approach while studying the ecol-
343 ogy of plague, which—like anthrax and many other bacterial pathogens—circulates easily
344 among fleas, rodents, other wildlife, humans, and the environment as one interconnected
345 system⁹⁰.

347 Developing a better understanding of plague in well-studied systems like the American

348 West will help develop a broader picture of its ecology. At present, all global maps of plague
349 foci have been compiled from expert knowledge; modeled products in the English language
350 are limited to the western United States, China, and Africa (see Extended Data Table 2).
351 In part, this reflects the challenges of sharing, aggregating, and consolidating surveillance
352 data. It may also likely reflects concerns about model transferability, given that the complex
353 multispecies dynamics of plague reservoirs differ greatly across ecosystems and continents.
354 However, other pathogens with regional host communities and complex environmental
355 persistence have been globally mapped through multinational coordination³⁸, and the same
356 synthesis is possible for plague. In the more immediate term, our model also strongly
357 suggests that wildlife reservoirs extend up to both national borders, and could plausibly
358 extend beyond them (recently confirmed for the northern border⁹¹), but the official World
359 Health Organization map of plague (last updated 2016) includes no reservoirs in Mexico
360 or Canada. Collaborating with national surveillance infrastructure in both countries may
361 help resolve the boundaries of plague transmission more clearly, and reveal foci currently
362 overlooked by global monitoring efforts.

363 Beyond plague, our study highlights the opportunity for medical geographers to de-
364 velop new methods that are suited to a rapidly changing world. Here, we proposed two
365 methodological advances that build on existing best practices in infectious disease map-
366 ping. First, time-specific covariates allowed us to train machine learning models on nearly
367 a century's worth of data, improving precision compared to coarsely-averaged predictors,
368 and capturing the effects of environmental change. If this approach is integrated with oth-
369 ers at finer temporal scales, such as those that consider seasonal aspects of transmission
370 or spillover risk^{19,20}, this could begin to set the foundation for an early warning system.
371 Second, the use of random intercepts to remove data and detection biases, such as the
372 serology method change in our data sample in 2011, is an important step towards testing
373 climate change impacts using continuous-time data (similar to how econometric approaches
374 resolve these problems in similar spatiotemporal analyses⁷). We propose that when this
375 approach can be taken, it may be used as a first principles method for detecting the signal
376 of environmental change in species' habitats. This could be a particularly important step
377 towards synthesizing the impacts of climate change on the shifting presence and absence
378 of disease data, especially in cases where prevalence and incidence data are lacking and
379 panel regression approaches cannot be applied. However, this work will still need to be
380 followed by proper "attribution" work that compares predicted patterns to counterfactual
381 scenarios without climate change; at present, all we can conclude with certainty is that
382 weather conditions have changed in a way that trends favorably for plague risk.

383 Our study also points to a number of gaps in our understanding of environmental change
384 (and consequently, potential methodological limitations). The PRISM data offers a fairly
385 comprehensive view of the recent climate in the United States, and allowed us to identify
386 the role of temperature and precipitation in plague transmission. However, we held both
387 soil and rodent predictor variables constant, and neither are stationary in reality. Soil has
388 changed over the last century due to a combination of climate change and land use change,

389 and unfortunately time-specific covariates are unavailable; in many cases, our soil layers had
390 to be generated custom to this study, and for the rest of the world these data are even more
391 sparse. Similarly, evidence is strong that most terrestrial species have responded to recent
392 climate change by undergoing range shifts, especially along elevational gradients. If rodents
393 have undergone range shifts, they may have encountered novel vector communities, and the
394 relationship between richness and transmission could change. Similarly, if elevation acts in
395 our models as a proxy for specific rodent-flea assemblages, range shifts could decouple the
396 observed relationships between elevation and transmission. As other studies have pointed
397 out, these challenges highlight the need to begin integrating zoonotic surveillance and
398 biodiversity monitoring¹.

399 **Methods**

400 Despite recent interest in modeling the distribution of major infectious diseases^{92,10}, there
401 is no definitive global map of plague reservoirs. All existing global plague maps have been
402 derived from expert opinion⁹³; all modeled products so far have been produced for national
403 or continental scales (see Extended Data Table 2). Plague ecology is regionally variable
404 enough that this patchwork approach has the advantage of being tailored to relevant local
405 predictors. However, the mix of modeling methods, variables, and spatiotemporal scales
406 makes it nearly impossible to compare these models and develop any consensus on the
407 biological or geological factors that determine where plague reservoirs can exist, and where
408 not. In this study, we adapt predictors that have previously worked in other similar work
409 on plague, and develop novel models of spatiotemporal risk patterns in the western United
410 States based on human and wildlife data spanning 1950 to 2017.

411 **Data**

412 Our study is designed around two independently-collected datasets, with only a small
413 amount of temporal overlap. Together, they provide as comprehensive a picture of plague
414 in the United States as possible.

415 **Human case data (1950-2005)**

416 Human cases of plague occur sporadically but consistently in the Western United States,
417 driven partially by exposure to infected cats and dogs that have acquired the infection out-
418 side of the home. The vast majority of cases are bubonic, though a handful of pneumonic
419 and septicemic cases occur. Confirmed plague cases are mandatorily reported to the U.S.
420 Centers for Disease Control and Prevention (CDC) Emergency Operations Center. CDC
421 surveillance data is actively maintained on plague, and has been previously published in
422 summary form as county totals⁵¹. We re-used these data, which have been anonymized
423 by previous researchers, and had case geolocations aggregated to county totals. To geo-
424 reference them, we randomly sample a number of locations within each county equivalent
425 to annual case totals. A total of 860 plague cases are recorded over the interval, with an
426 average of 7.7 cases per year, across 490 counties in the American west.

427 **Wildlife serology data (2000-2017)**

428 Wild animals are routinely exposed to *Y. pestis* in endemic regions, including the United
429 States. Infection leads to substantial morbidity and mortality in some species (e.g., prairie
430 dogs), but other species (e.g., coyotes) readily survive infection, with antibodies to *Y.*
431 *pestis* being the only indication of exposure. This is especially true for predators, which
432 can be exposed through consumption of plague-positive rodents or through bites from
433 plague-positive fleas. These predator species do not necessarily play a direct role in plague

434 transmission and dynamics, but instead act as sentinels of plague activity on the land-
435 scape^{94,95}. Correspondingly, the USDA National Wildlife Disease Program tests wildlife
436 for evidence of plague exposure throughout much of the western U.S. Testing was con-
437 ducted using a hemagglutination assay⁹⁶ at the Centers for Disease Control and Prevention
438 until 2011. A majority of samples collected after 2011 were tested using a bead-based flow
439 cytometric assay with a lower limit of detection⁹⁷. In total, the version of the dataset we
440 used spanned February 13, 2000 to January 29, 2018, with a total of 41,010 records, in-
441 cluding 5,043 animals that tested positive. Of those records, the vast majority are coyotes
442 (32,825 animals including 4,812 that tested positive).

443 **Environmental covariates**

444 The transmission ecology of plague shares features with both vector-borne systems (e.g.,
445 malaria or dengue fever) and soil-borne pathogens (e.g., anthrax or melioidosis). The
446 predictors we have chosen here are informed by predictors that have performed well for
447 other plague mapping projects (see Extended Data Table 2), and were all expected to be
448 informative as drivers of host ecology, vector competence and/or soil persistence.

449 Most studies that map infectious diseases with machine learning methods (i.e. eco-
450 logical niche models) use long-term climate averages, paired with occurrence data that
451 sometimes span decades of unstable environmental conditions. In contrast, we used time-
452 specific climate data paired with—and extracted for—the year of each data point in the
453 occurrence data. This allowed us to make yearly spatial predictions of the distribution
454 of plague risk, and consider the extent of transmission risk as a dynamic process rather
455 than a static surface. We held non-climate predictors constant, assuming them to either be
456 invariant (elevation) or long term averages (soil and rodent richness); in a more advanced
457 retrospective, it might be possible to reconstruct the impacts of land use change by adding
458 yearly resolution to these covariates, but these data do not currently exist.

459 **Climate**

460 We derived all climate data (1950–2017) from PRISM, a historical reconstruction of cli-
461 mate in the continental United States, derived from a mix of weather station data and
462 climatologically-aided interpolation.⁹⁸ From the PRISM dataset, we used cumulative an-
463 nual precipitation and annual mean, minimum, and maximum temperatures. We also
464 generated two “anomaly” variables, given on a pixel-by-pixel basis as the difference be-
465 tween the annual value and the long-term average, divided by the variance. These data
466 were downloaded in 2.5 arcminute grids (~4.5 km at the equator), which was used as the
467 standard resolution for the rest of the project.

468 **Soil**

469 We assembled a set of seven predictor layers for soil persistence of plague that were informed
470 by both laboratory experiments on plague transmission, and previous efforts mapping soil-
471 borne pathogens like anthrax (*Bacillus anthracis*) and botulism (*Clostridium botulinum*).
472 We aimed to develop a cohesive set of predictors characterizing the C layer (~1m depth);
473 rodent burrows in the American west can go up to two meters deep in the soil, but macronutri-
474 ent data is limited at this depth. We used the most recent version of the ISRIC SoilGrids
475 global dataset at 250 meter resolution⁹⁹, and selected gridded layers of soil pH, cation ex-
476 change capacity (base saturation), and the concentration of sand, clay, and organic content
477 in the top 60-100 mm layer of soil. Sodium, calcium, and iron concentrations were derived
478 from a national survey of soil geochemical properties, published in raw form as USGS data
479 series 801.^{100,101} We extracted all point samples of mineral concentration in the C horizon,
480 given in weight percent, and then developed a rasterized layer for these macronutrients by
481 kriging the point data, using the `autoKrig` function in the `automap` package.

482 **Additional covariates**

483 Rodent species richness was derived by stacking species IUCN expert range maps for the
484 Rodentia, and rasterizing the richness layer using the `fasterize` package. Elevation data
485 was scraped using the `elevatr` package in R, which pulls gridded elevation data from
486 the AWS Open Data Terrain Tiles. We pulled elevation data at resolution “6”, which
487 returns elevation rasters in 2,446 meter squared grids at the equator (~1.3 arcminutes),
488 and aggregated to the native resolution of the other grids.

489 **Modeling**

490 Dozens of statistical methods have been applied to species distribution modeling in the
491 past few decades, with a wide range of performance.¹⁰² Over the past few years, clas-
492 sification and regression tree methods (CART) – including random forests and boosted
493 regression trees – have become especially popular for mapping the geographic distribution
494 of infectious diseases^{103,38,104,105,106,107}. Here, we use a fairly new method, Bayesian addi-
495 tive regression trees (BARTs), implemented with the R package `embarcadero` as a species
496 distribution modeling wrapper for the `dbarts` package.⁵³ BART is a powerful new method
497 with growing application in computer science, and often performs comparably to other
498 CART methods like random forests and boosted regression trees.¹⁰⁸ In the `embarcadero`
499 implementation, BARTs have several unique features that make them a powerful tool for
500 disease mapping, such as: model-free variable importance measures, and automated vari-
501 able selection; posterior distributions on predictions, as a measure of uncertainty; posterior
502 distributions on partial dependence plots; two-dimensional and spatially-projected partial
503 dependence plots; and various extensions, including random intercept models.

504 Like other CART methods, BART makes predictions by splitting predictor variables
 505 with a set of nested decision rules (“trees”) that assign estimated values to terminal nodes
 506 (“leaves”). Predictions are generated based on a sum-of-trees model, where a set of n trees
 507 with leaves $(T_1, M_1), \dots, (T_n, M_n)$ each make predictions $g(\cdot)$ that are added together, for a
 508 total estimate:

$$\hat{Y} = f(X) = \sum_{j=1}^m g(X; T_j, M_j) + \epsilon; \quad \epsilon \sim \mathcal{N}(0, \sigma^2)$$

509 For logistic classification problems (like species distribution modeling), BART uses a logit
 510 link function:

$$\hat{Y} = f(X) = \Phi \left[\sum_{j=1}^m g(X; T_j, M_j) \right]$$

511 where Φ is the standard normal cumulative distribution. An initial set of n trees is fit,
 512 and then altered in an MCMC process based on a set of random changes to the sum-of-
 513 trees model (e.g., new splits added, levels rearranged, or leaves pruned). An initial burn-in
 514 period is discarded, and then a set of posterior draws of f^* create the posterior distribution
 515 for $p(f|y) \equiv p(\text{trees}|\text{data})$.

516 BART is easily implemented out-of-the-box, even with a full Bayesian MCMC component.
 517 Three priors control the ways decision trees change: the probability each variable is
 518 drawn for a split, the probability of splitting values tested, and the probability a tree stops
 519 at a certain depth. In the simplest form, the first two can be set as uniform distributions,
 520 while the latter is usually set as a negative power distribution; they can also be adjusted us-
 521 ing a full cross-validation approach. This is handled automatically in the `dbarts` package,
 522 for which `embarcadero` is a wrapper. More advanced implementations with complex prior
 523 design are sometimes appropriate: for example, a Dirichlet distribution on the variable
 524 importance prior can help identify informative predictors in high dimensionality datasets
 525 (dozens or hundreds of covariates). However, in our case, we had confidence all variables
 526 were biologically plausible based on expert opinion.

527 The base models

528 We ran two separate baseline models, the first using human data from 1950 to 2005, and
 529 the second using the wildlife data from 2000 to 2017. For the human model, we used the
 530 number of cases recorded each year in each county to generate a set of random georeferenced
 531 pseudopresence points. We then generated seven pseudoabsence points in each year to
 532 create a roughly balanced design, for a total of $n = 430$ pseudopresence points and $n =$
 533 392 pseudoabsence points. For the wildlife model, we balanced the design by subsampling
 534 seronegative animals in equal number to seropositive ones, for a final $n = 5,002$ true
 535 presence points and $n = 4,759$ true absence points.

536 Both models were run with the full predictor set, followed an automated variable set
537 reduction procedure implemented in `embarcadero` that formalizes the recommendations of
538 Chipman *et al.* (2010).¹⁰⁸ In BART, variable importance is “model-free,” measured as the
539 number of splitting rules involving a given variable (but incorporating no information on
540 the proportional effect on the outcome variable, or proportional improvement of the model
541 predictions). In models with fewer trees (small n), informative variables tend to be selected
542 more often, while uninformative variables are selected rarely or drop out entirely. This
543 property of BART establishes a rubric that can be used to identify an informative variable
544 set. First, an initial model is fit with all variables 100 times each for six different settings
545 of ensemble size ($n = 10, 20, 50, 100, 150$, and 200 trees). Plotting the average importance
546 of variables at each level offers a qualitative diagnostic of how informative each predictor
547 is. Next, an initial set of 200 models with $n = 10$ trees are run, and variable importance is
548 recorded and averaged across models. Models are run again (200 times) without the least
549 informative variable from the first fit, and this is performed iteratively until only three
550 variables remain; the variable set with the lowest average model root mean square error
551 (RMSE), and therefore highest accuracy on the training data, is selected. Finally, we plot
552 variable importances (including standard deviations based on model permutations).

553 Final models were run with the reduced variable set, with recommended BART model
554 settings (200 trees, 1000 posterior draws with a burn-in of 100 draws) and hyperparameters
555 (power = 2.0, base = 0.95 for the tree regularization prior, which limits tree depth). We
556 then used the `retune` function in `embarcadero` to run a full cross-validation panel on the
557 three prior parameters. `retune` runs a full cross-validation across the k hyperprior (values
558 of 1, 2, and 3), the base parameter (0.75 to 0.95 in increments of 0.05), and the exponent
559 parameter (1.5 to 2 in increments of 0.1), and returns the model with the parameter
560 combination that generates the minimum root mean squared error.

561 For the wildlife model, the final variable set included: temperature mean, maximum,
562 and anomaly; rodent richness; elevation; and five soil traits (calcium, sodium, iron, clay,
563 and sand). The model validated well on training data (AUC = 0.836). For the human
564 model, the final variable set included a similar subset: precipitation anomaly; temperature
565 mean, maximum, and anomaly; rodent richness; elevation; and four soil traits (sodium,
566 iron, clay, and calcium). The model also validated well on training data (AUC = 0.909).

567 Alternate formulations

568 As a final check of model performance, we ran a separate model with the same predictor
569 sets that withheld the years 2000–2005 from both. On the test dataset for humans ($n =$
570 64), the model performed very well by the standards of external cross-validation (AUC =
571 0.820); on the test data for wildlife ($n = 796$), the model also performed well (AUC =
572 0.775). This indicated that both models were performing adequately.

573 We also recognize that model design can have a substantial effect on machine learning
574 performance, and the downstream biological inference made by using ecological niche

models. Given that BART is a relatively new method, it has been comparatively underexplored in this regard, and so a standard panel of “best practices” has not yet been recommended in the literature. However, for transparency about model uncertainty and the influence of subjective decisions on model outputs, we produced four major alternate formulations. First, we produced models that included all variables, rather than using the variable set reduction procedure, for both the human data (Extended Data Figures 20,21, 22) and wildlife data (Extended Data Figures 23,24, 25). We additionally considered two alternate formulations of the wildlife model. In the first, we used pseudoabsences instead of the true absences available in the data (Extended Data Figure 26). Though this increased model AUC (0.929), and allowed slightly different balancing of the data, it lead to visually-apparent overfitting. Finally, we ran an alternate model only using the coyote data in the NWDP dataset, which also performed adequately (AUC = 0.826; Extended Data Figure 27). Both models were ultimately not selected because they left available, biologically-meaningful data unused, and both produced predictions that were slightly less congruous with the human model.

590 **Prediction, delineating foci, and measuring change**

Although the models were trained over different intervals, the continuous and standardized set of predictors allowed cross-prediction over the entire extent of the study (1950–2017). For each layer of annual prediction, we thresholded suitability based on a model-specific threshold chosen to maximize the true skill statistic on the test data. We mapped areas of “unstable foci” as any region with at least one year of suitability, and “stable foci” as any region suitable in every year over the 70-year interval. This allowed us to compare long-term spatial patterns between the two models.

598 **Random effects models for interannual variation**

Prevalence changes year-to-year in both the data and modeled landscapes, but detecting the signal of climate change in that fluctuation can be challenging. There are several reasons prevalence could vary across years: (1) incidence is stochastic but temporally autocorrelated; (2) normal climatic variability (e.g. the Pacific Decadal Oscillation) or other socioecological trends (e.g., rising human populations) might also contribute to interannual variation, including non-linear trends over time; (3) anthropogenic climate change is directly driving changes in plague risk, or indirectly changing the ecology of the involved species; (4) sampling effort varies between years (for wildlife); or (5) detection rates could change between years, due to testing or surveillance. The last of these is particularly relevant as a possible confounder, given that wildlife diagnostics changed in 2011. A positive trend in plague risk might be generated by increased climatic suitability for plague transmission, but could also be generated by a consistent increase in plague detection due to improved diagnostics and increased sampling effort, loosely colinear with warming temper-

612 atures on the scale of 20 to 70 years.

613 We propose a new method that uses machine learning approaches (i.e., ecological niche
614 models) to detect the signal of environmental change while adjusting for confounders at
615 a high level. The approach is loosely modeled off the ideas underlying econometric ap-
616 proaches to climate change detection and attribution, which usually use fixed effects panel
617 regression to control for spatiotemporal confounders in climatic signal. By attributing as
618 much variance as possible to spatial, temporal, and other confounders, and then identifying
619 climatic signal in the remaining variance, these approaches can pinpoint the signal of envi-
620 ronmental change with a high degree of confidence. So far, no analog to these approaches
621 exists for ecological niche models. Only a handful of studies have even added temporal
622 heterogeneity to ENMs; so far, we know of none that have also independently controlled
623 for interannual variation in detection, sampling effort, or species prevalence.

624 A solution to temporal confounders is particularly needed in this study, given the chal-
625 lenges of the time-specific approach. In default settings, BART predictions converge on
626 observed prevalence, i.e., $\hat{Y} = E[Y] = P(Y = 1)$. For this reason, we balance the presences
627 and absences, so that the model is as close to $\hat{Y} \sim \mathcal{N}(0.5, \sigma^2)$ as possible. This produces a
628 unique challenge for time-dependent modeling. Presences are distributed unevenly across
629 years, and consequently, so is positivity. In the human model, this arises artificially, be-
630 cause pseudocabsences are generated evenly across years. We chose this approach to avoid
631 over-representing years with more cases in the data, which would introduce an additional
632 colinearity, but as a result relative prevalence varies substantially. This bias also affects
633 the wildlife model more organically; although the number of points per year varies inde-
634 pendently of test positivity ($\text{cor} = 0.113$, $p = 0.687$), because most sampling is passive,
635 there is still a wide range in annual prevalence (28% in 2006 versus 80% in 2017, both
636 years with several hundred records), with a net trend towards higher positivity over time.
637 Because prevalence varies between years in both models, the resulting colinearity with
638 environmental change could confound the detection of meaningful signals.

639 Inspired by the econometric approach, we propose a use case for the *random intercept*
640 *BART model* (riBART), which has recently been proposed as an extension of the method
641 for clustered outcomes. The approach adds a random intercept term to the model (separat-
642 from the tree-fitting process) based on the identified K clusters

$$\hat{Y} = f(X) = \Phi \left[\sum_{j=1}^m g(X_k; T_j, M_j) \right] + \alpha_k$$

643 where the random intercepts α_k ($k \in 1 : K$) are normally distributed around zero (i.e., the
644 K groups are assumed to have normally distributed, independent additive effects on the
645 outcome variable). The error structure of the random effects and the sum-of-trees model
646 are assumed independent. Here, we propose that the model can be fit as usual with a
647 random intercept for year, as a way of accounting for temporal heterogeneity as a possible
648 confounder. The yearly random intercept absorbs most of the interannual variation in

649 plague prevalence (i.e., the relative ratio of presences and absences in the data), such that
650 residual variation in prevalence should be roughly constant across years. Identifying the
651 climatic signal in this residual data, and then examining predicted prevalence (without
652 random intercepts) based only on environmental change, allows more confident statements
653 about how environmental change contributes to shifting disease risk.

654 We revisited the two main models, and used riBART to add an annual random intercept
655 to our model for each year, which we refer to throughout as the “detection” models. Fit-
656 ting climate-plague response curves after this detrending decouples the possible colinearity
657 between climate trends and coarse interannual signal in the data, which may be driven
658 by natural variation in prevalence or other confounders (e.g, the 2011 change in wildlife
659 testing protocols). We fit both detection models with a random intercept for year, plotted
660 the random effects, and predicted over the 70 year interval without the random effect in-
661 cluded. (All functionality to implement SDMs with riBART is available in `embarcadero`
662 as an updated release.)

663 **Detecting change over time**

664 To estimate trends of change over time, we fit a linear slope through each pixel-by-year.
665 Multiplying by 68 years, we were able to estimate total percent change in suitability since
666 1950 in a given pixel. We did not limit these to pixels with a significant trend, as any
667 frequentist significance test iterated over millions of pixels would be mostly meaningless.
668 We generated these maps for the two primary models and the two detection models (Figure
669 3), as well as (in the supplement) for mean temperature and precipitation (Extended Data
670 Figures 18,19).

671 **Data and code availability**

672 Human case data in this study is taken from previous studies and is available online for
673 researchers to reproduce our study. Wildlife data is available on formal request and
674 approval from the United States Department of Agriculture. All code is available at
675 github.com/cjcarlson/plague-wna.

676 **Figures**

Figure 1: Suitability for plague across all years (1950-2017), for humans (left) and wildlife (right). Top panels give mean suitability across all years; bottom panels show areas identified as suitable in no years, at least one, or all 68 years.

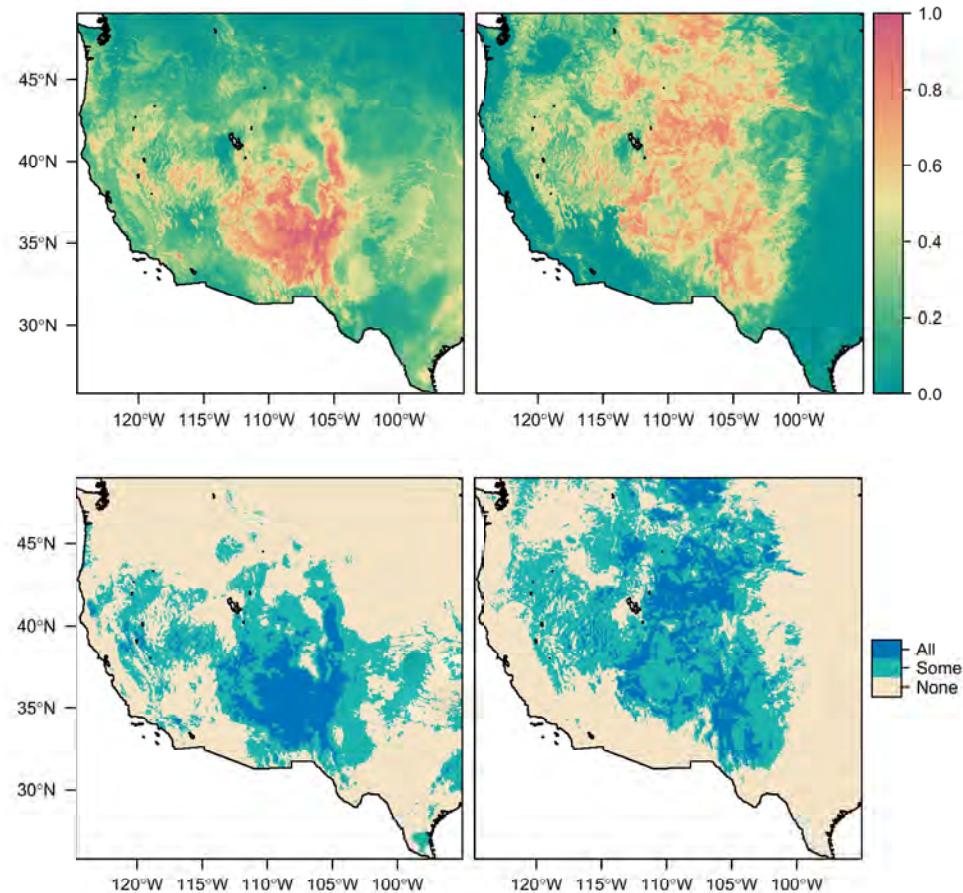


Figure 2: The wildlife model's predictions largely encompass the human model's predictions, except in southern Arizona and California (where predictions extend into other areas of suspected plague risk) and west Texas (too far east for plague reservoirs).

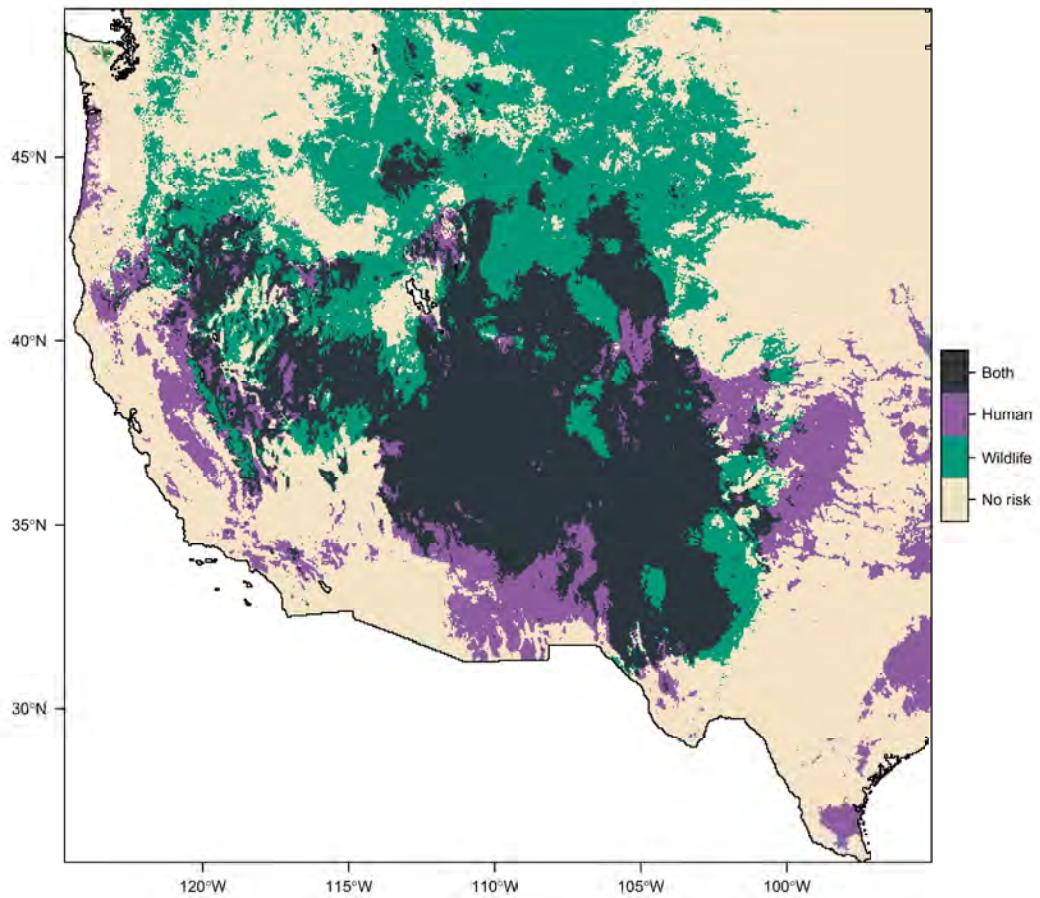


Figure 3: Total percent change in plague suitability, 1950 to present, in the human (top) and wildlife models (bottom), before (left) and after (right) adding random intercepts to control for interannual variation.

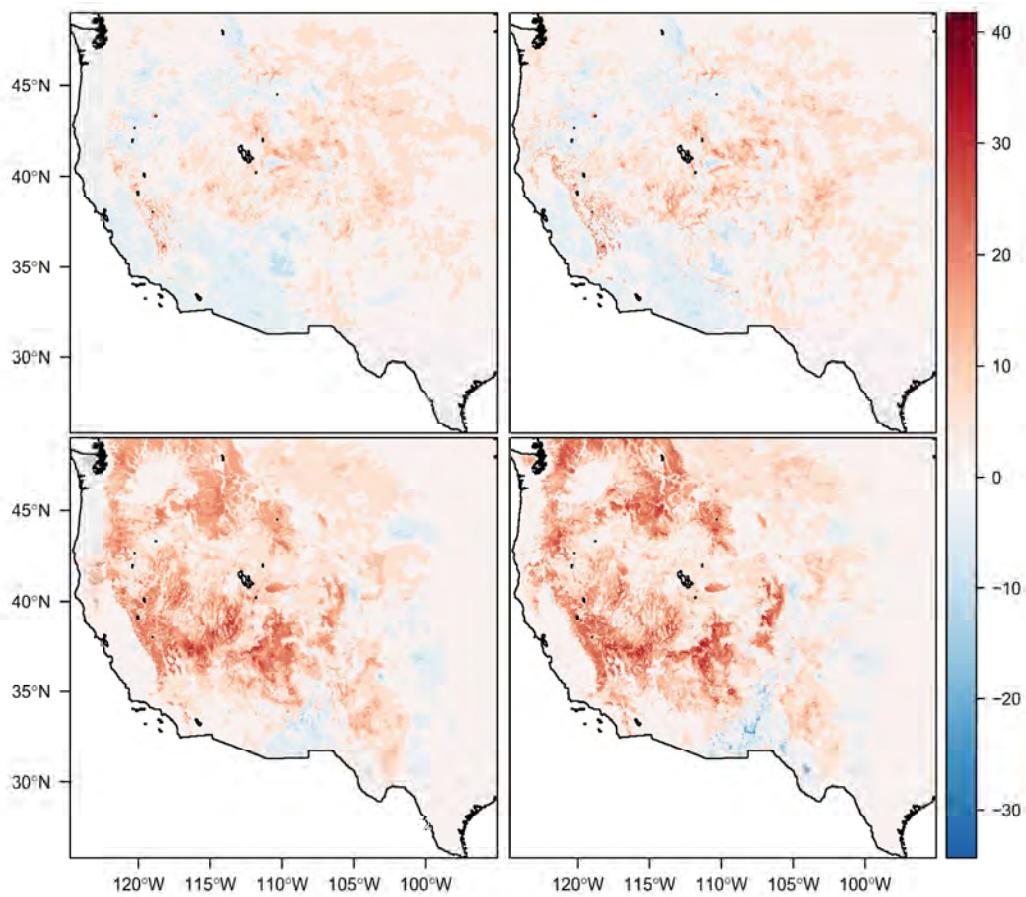
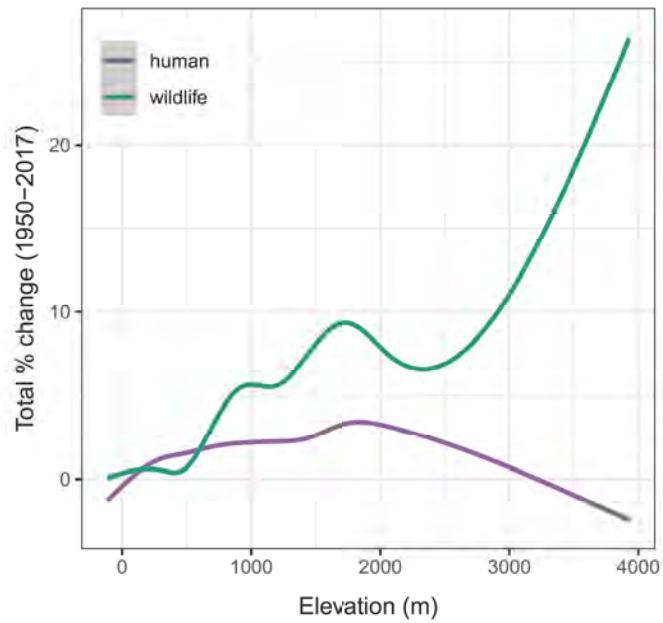
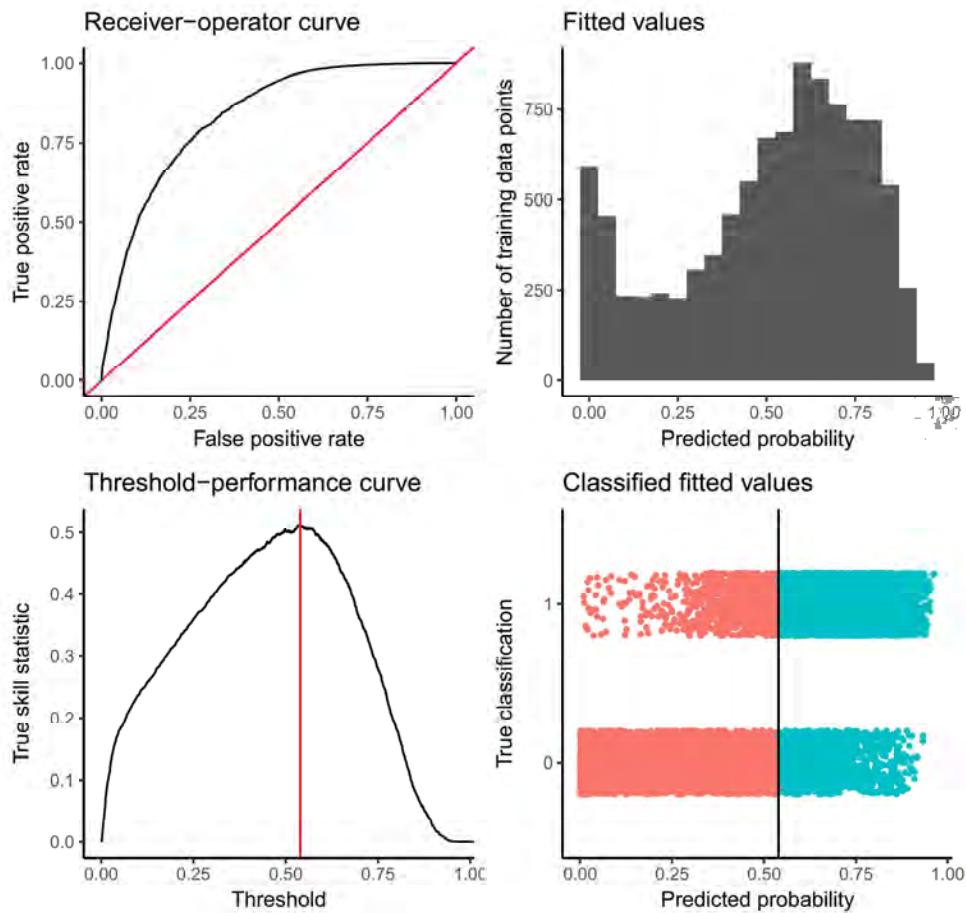


Figure 4: Environmental suitability for plague has increased substantially at high elevations for wildlife; risk of spillover has increased mildly at mid-elevations. Lines are given as generalized additive model smooth fits based on the detection models.

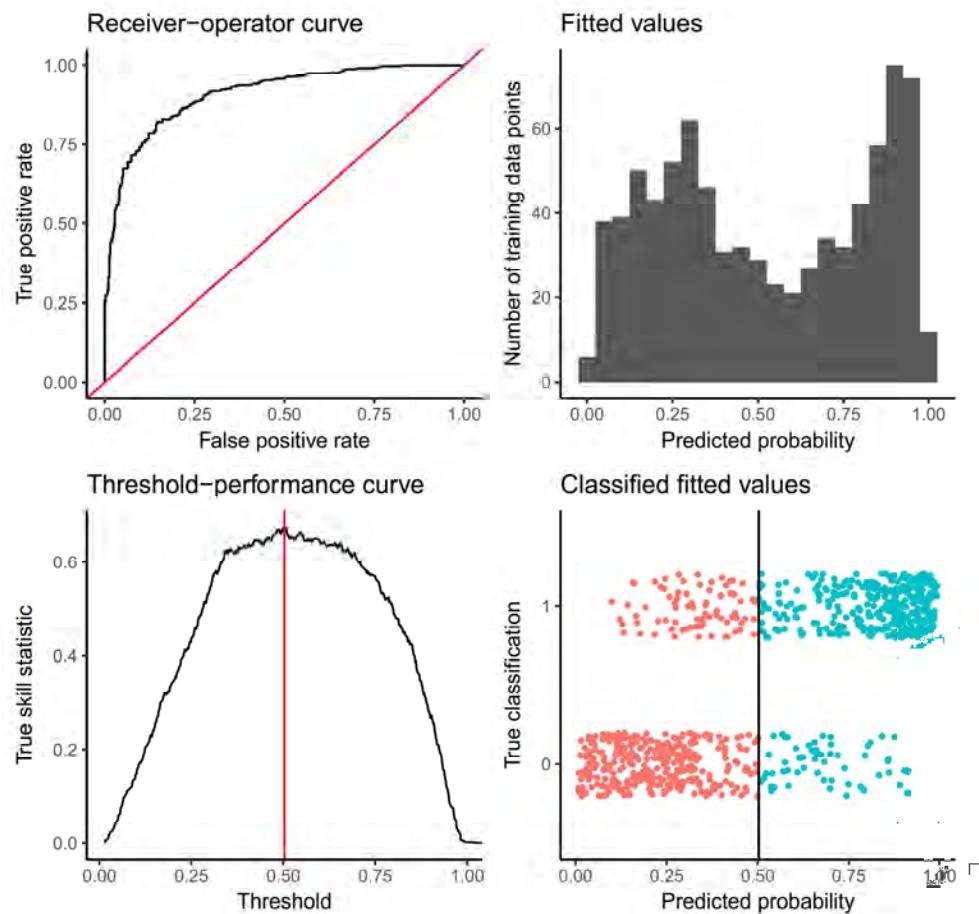


Supporting Information

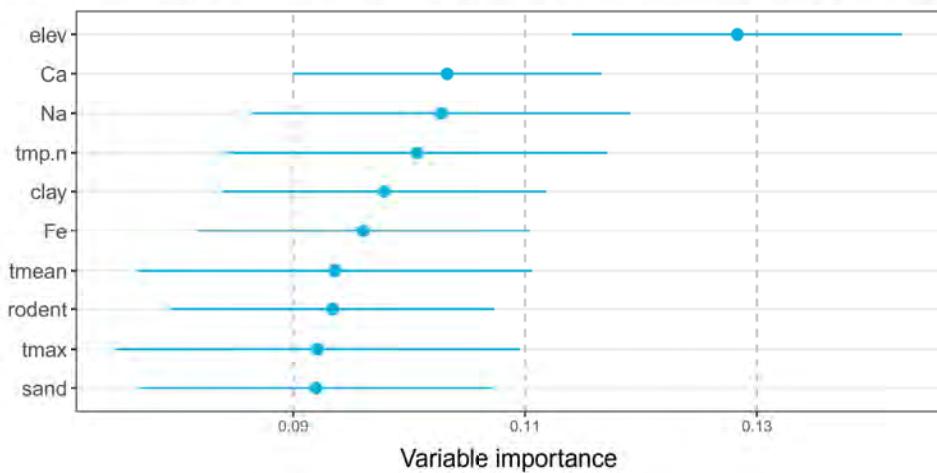
Extended Data Figure 1: Summary model diagnostics for the wildlife plague risk model.



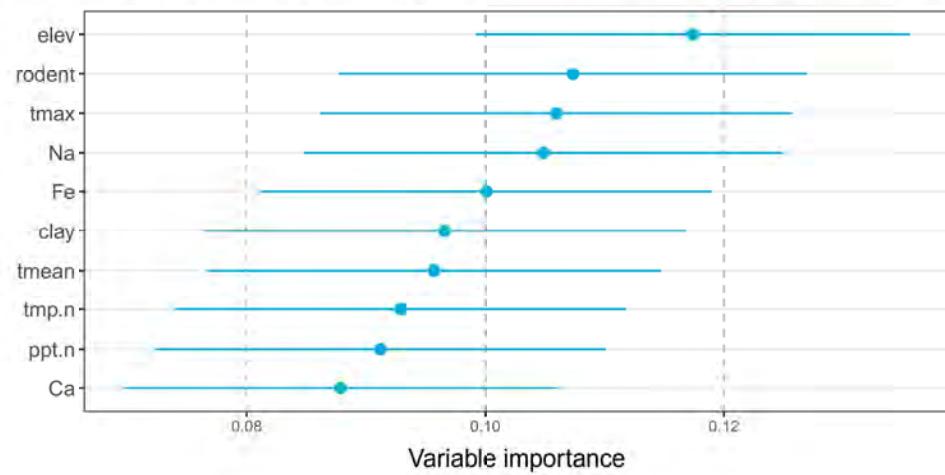
Extended Data Figure 2: Summary model diagnostics for the human plague risk model.



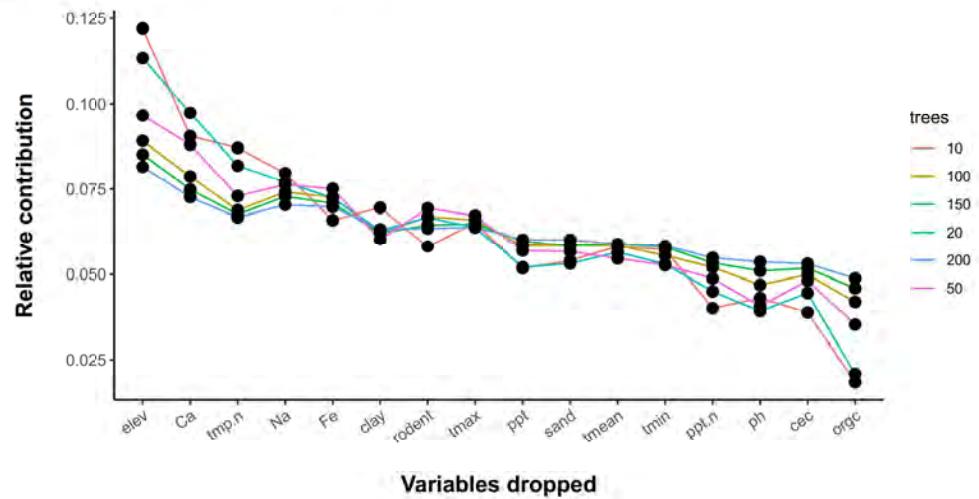
Extended Data Figure 3: Variable importance in the wildlife plague risk model.



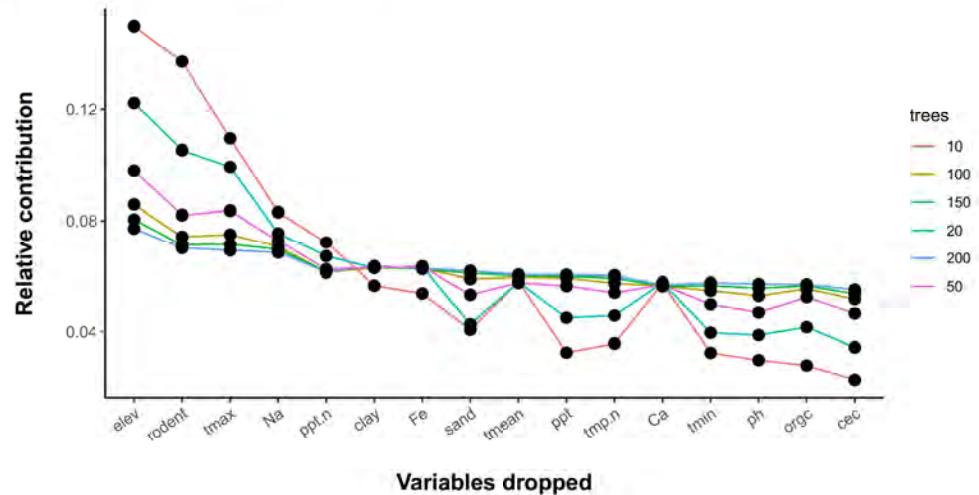
Extended Data Figure 4: Variable importance in the human plague risk model.



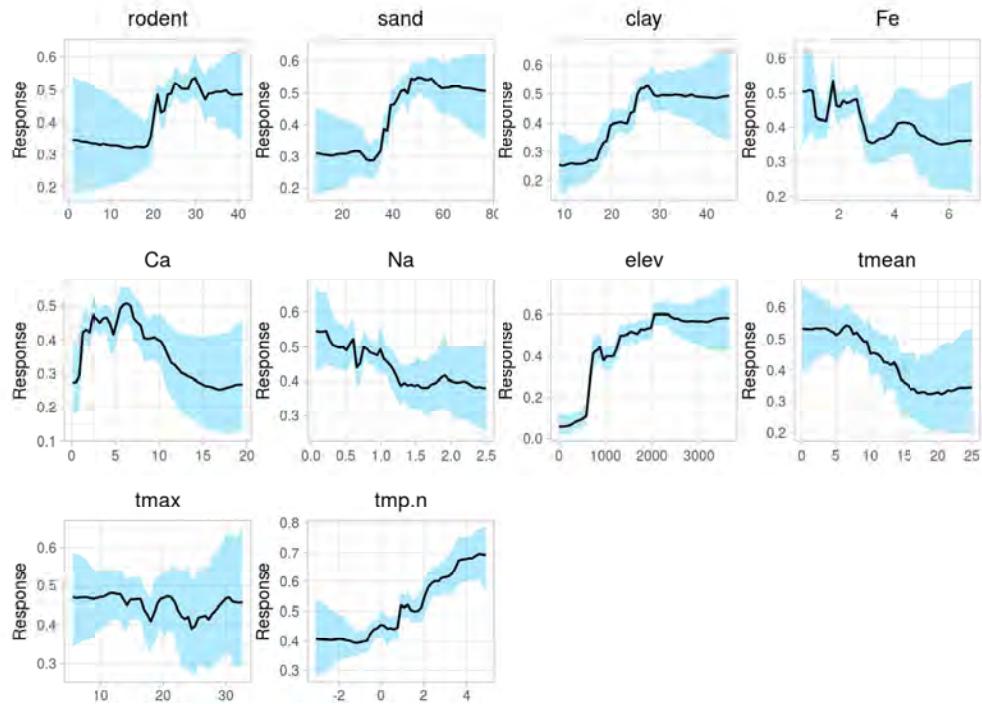
Extended Data Figure 5: The variable importance diagnostic for all variables considered for the wildlife plague risk model.



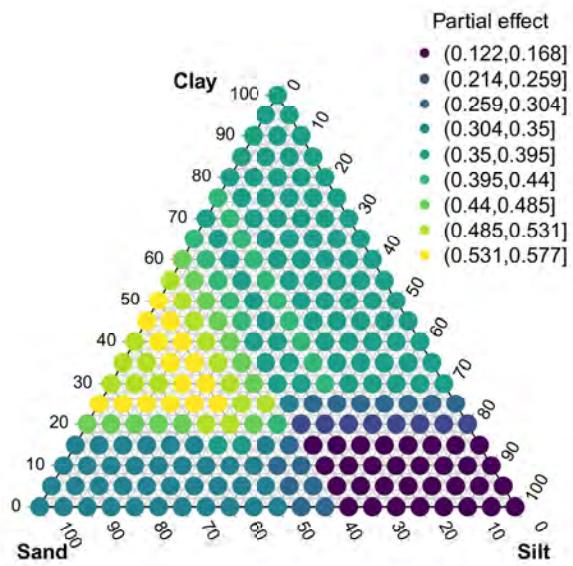
Extended Data Figure 6: The variable importance diagnostic for all variables considered for the human plague risk model.



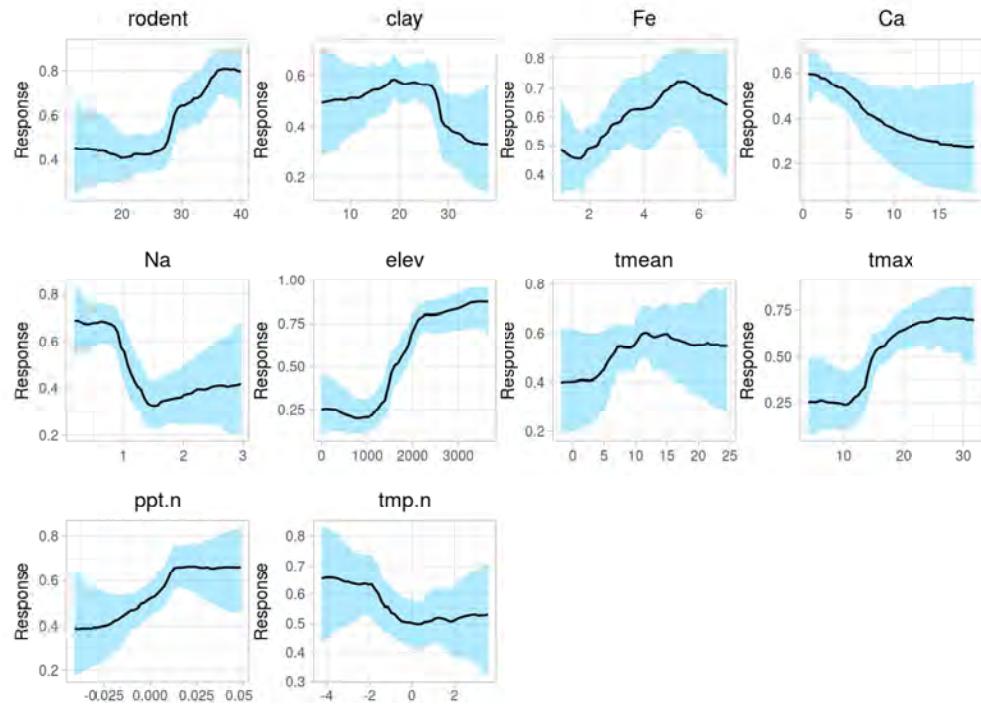
Extended Data Figure 7: Full partial dependence plots for the wildlife plague risk model.



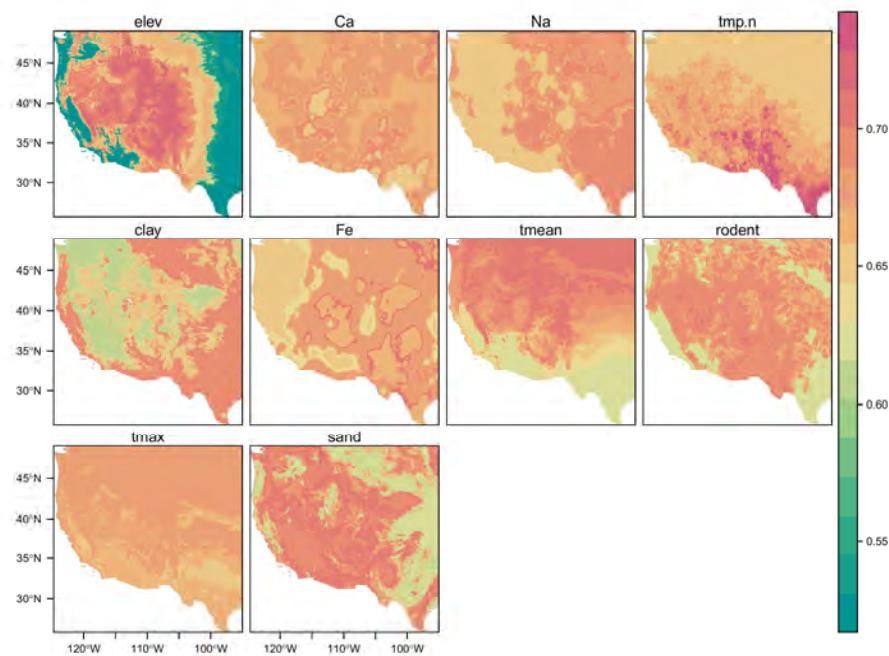
Extended Data Figure 8: Two-dimensional partial dependence plot for sand and clay in the wildlife model, projected onto the soil composition triangle as a ternary partial (tertial) plot.



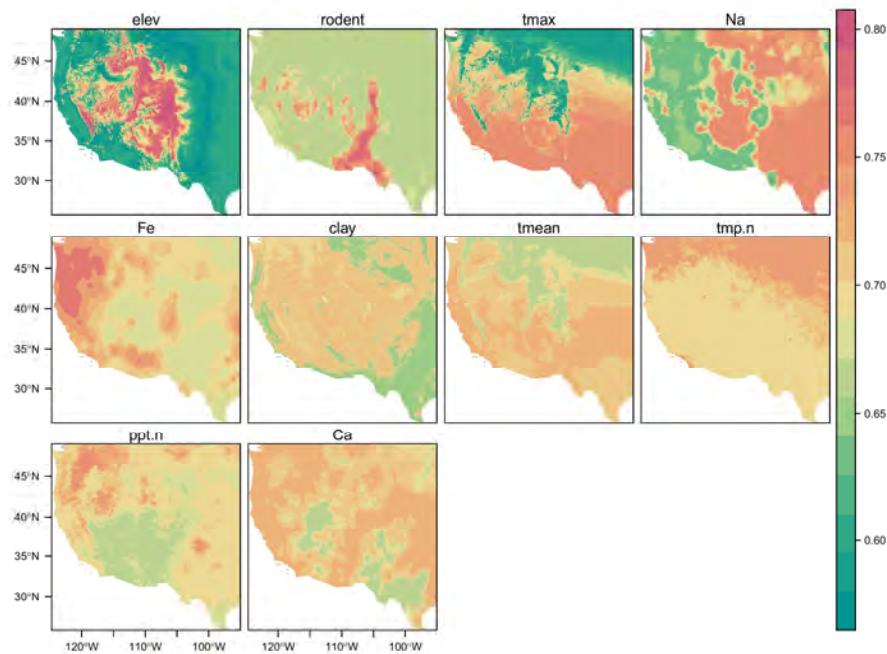
Extended Data Figure 9: Full partial dependence plots for the human plague risk model.



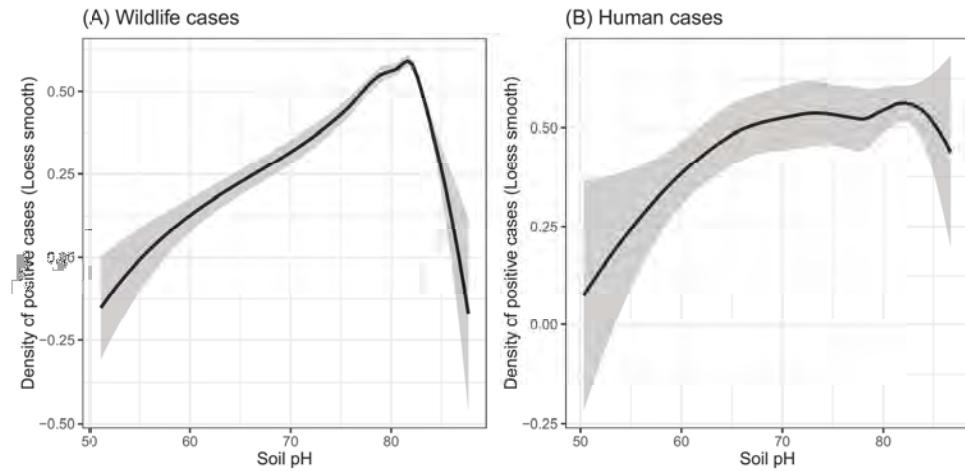
Extended Data Figure 10: Spatial partial (spartial) dependence plots for the wildlife plague risk model. Variables are given in descending order of model importance.



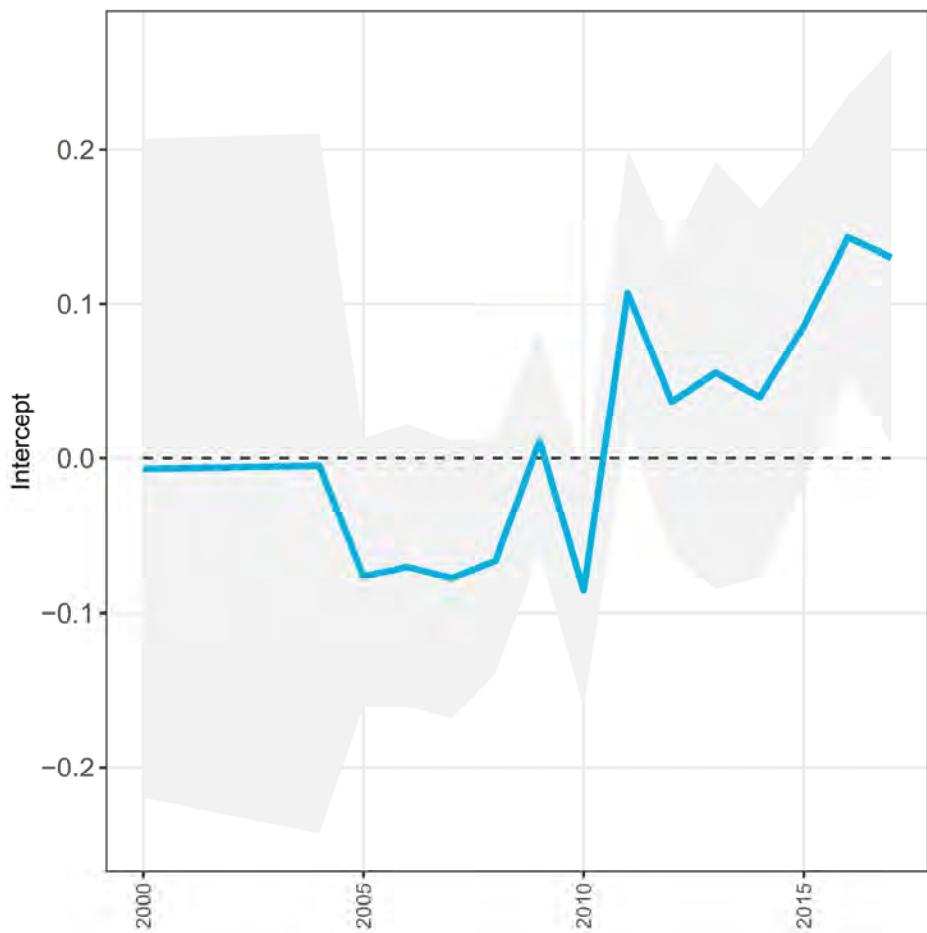
Extended Data Figure 11: Spatial partial (spartial) dependence plots for the human plague risk model. Variables are given in descending order of model importance.



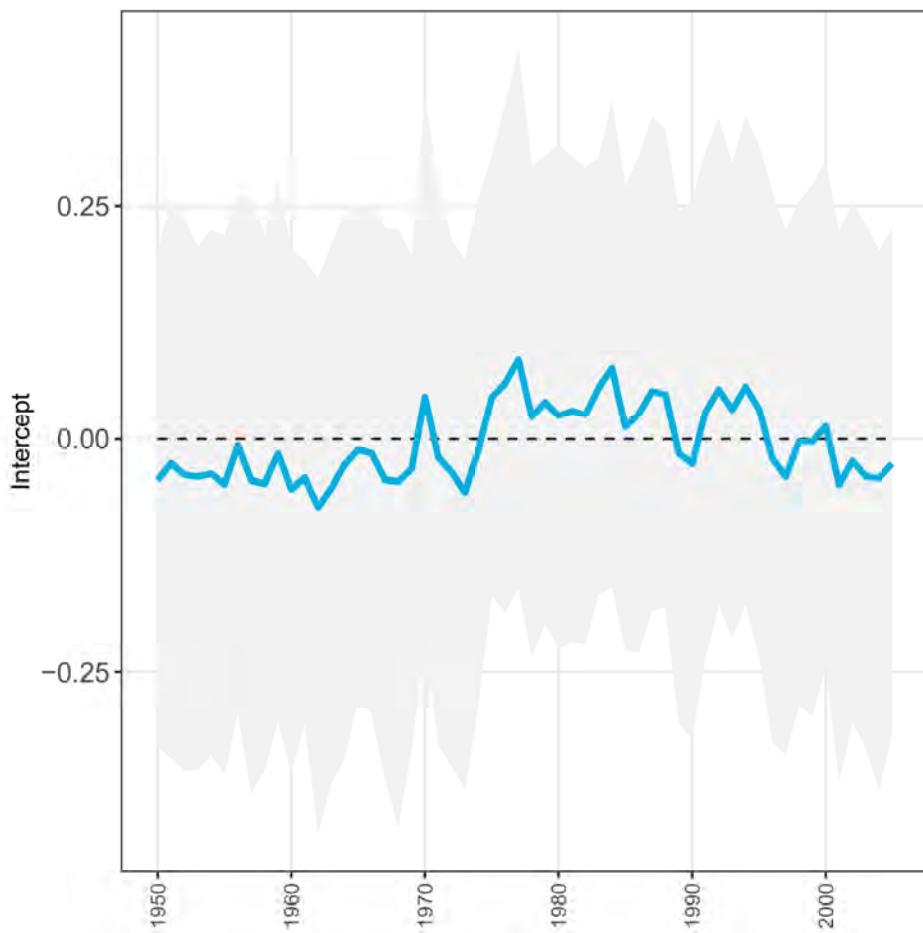
Extended Data Figure 12: Density of positive cases in the wildlife and human case data relative to recorded pH, modeled using a simple Loess smooth.



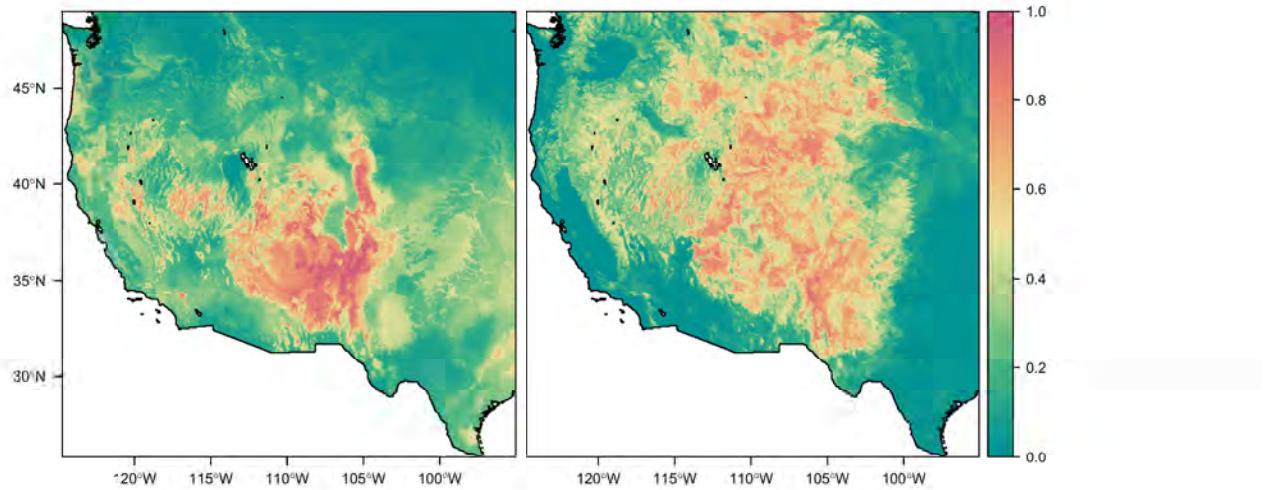
Extended Data Figure 13: Random intercept values for each year in the detection model for wildlife plague risk.



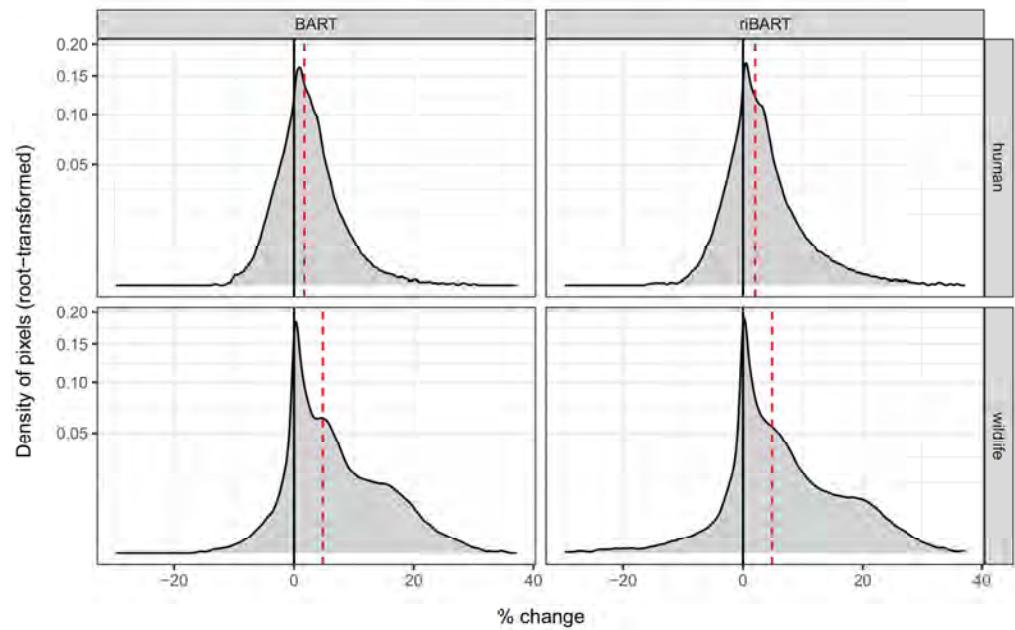
Extended Data Figure 14: Random intercept values for each year in the detection model for human plague risk.



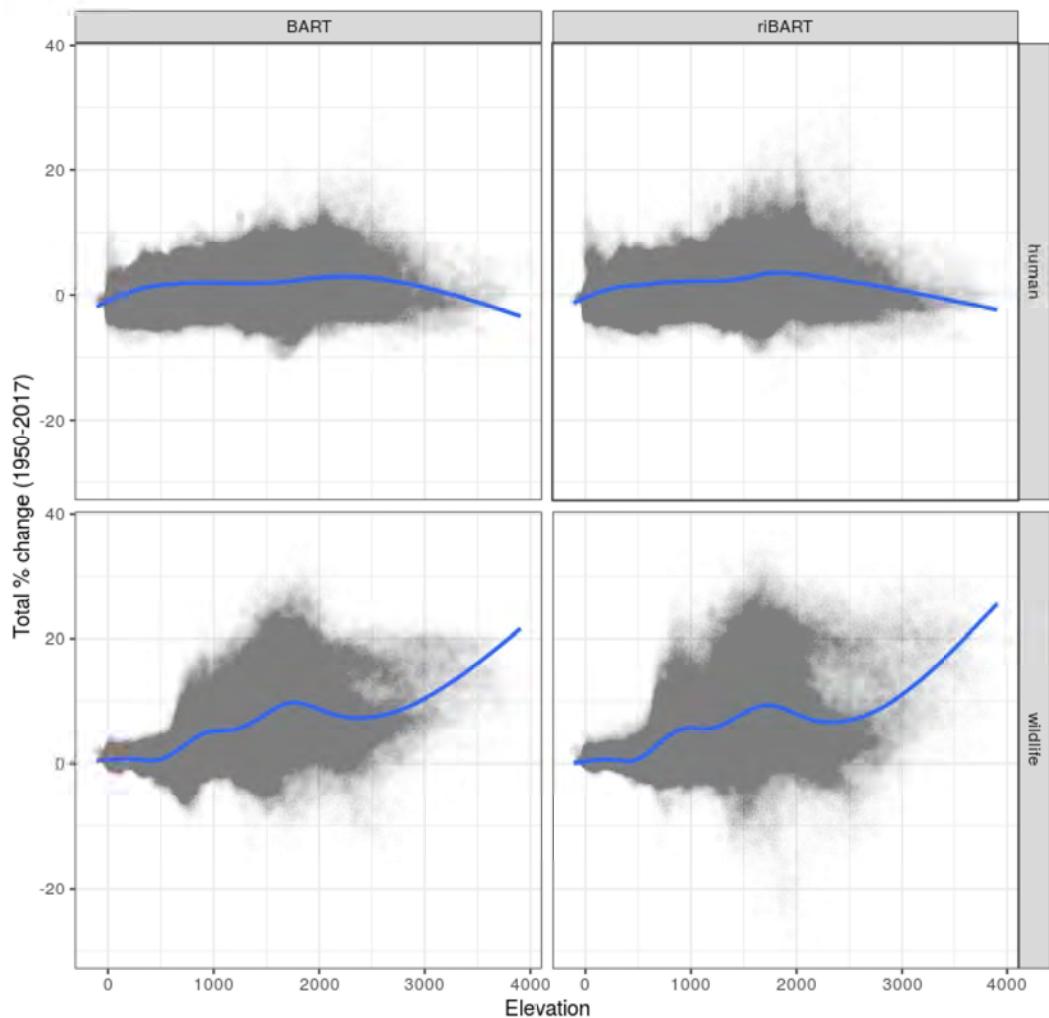
Extended Data Figure 15: Mean environmental suitability across years in the detection model (left: human; right: wildlife). Random intercept models have nearly identical predictions to baseline models; the two are nearly perfectly correlated for humans ($r = 0.992$) and wildlife ($r = 0.969$).



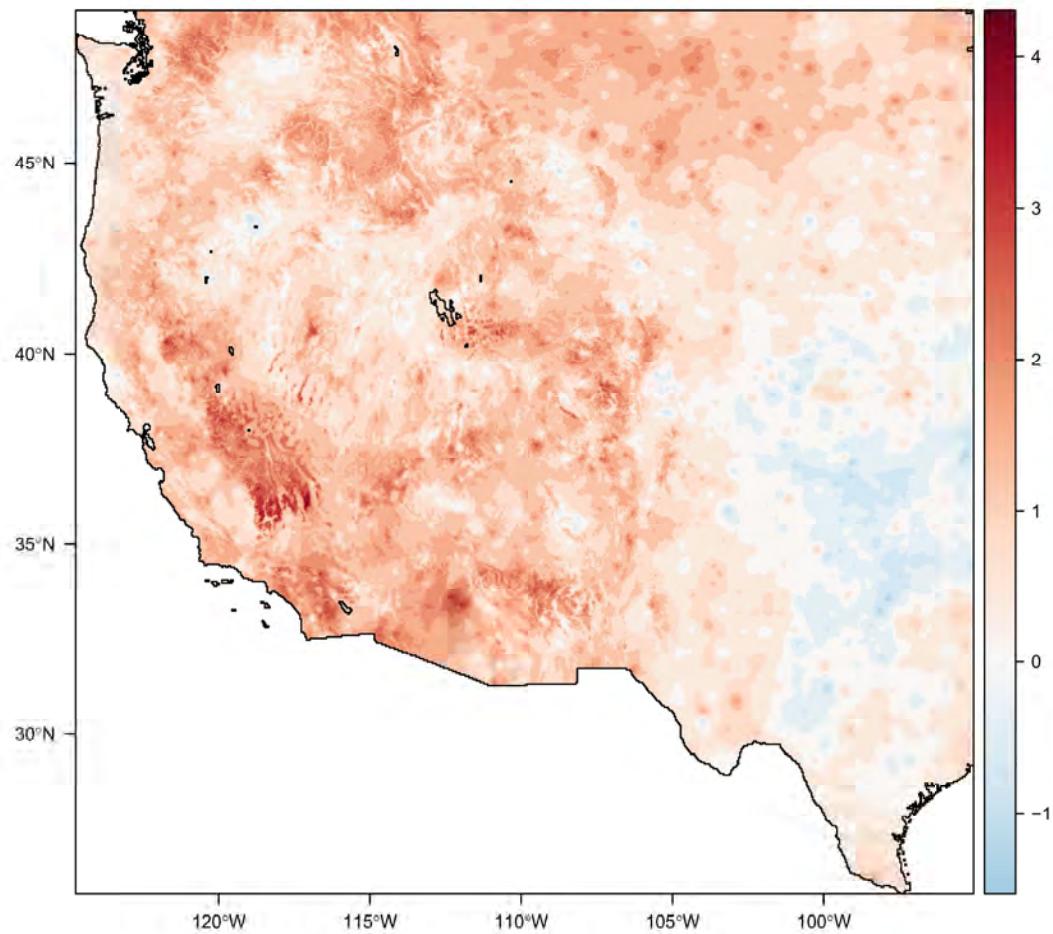
Extended Data Figure 16: Estimated percent change in plague suitability, 1950 to present, across pixels and the four main models. Red lines show the mean value.



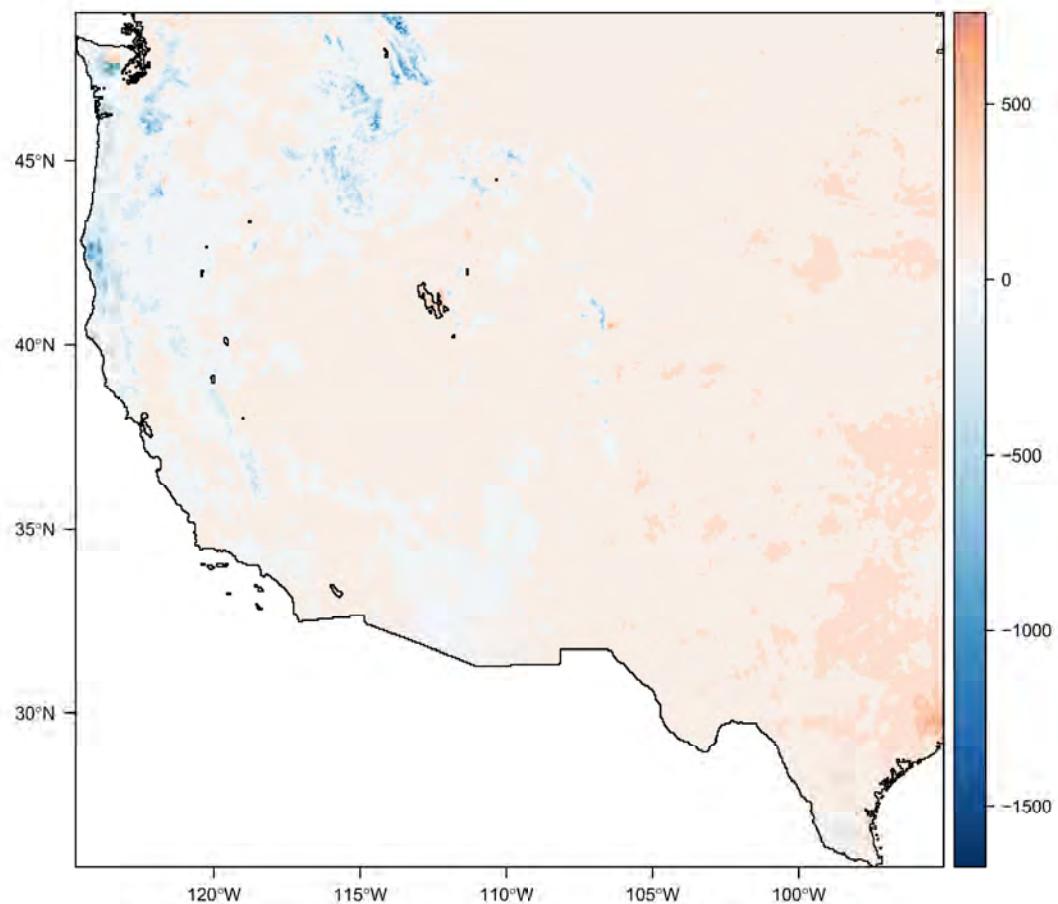
Extended Data Figure 17: Estimated percent change in plague suitability, 1950 to present, versus elevation.



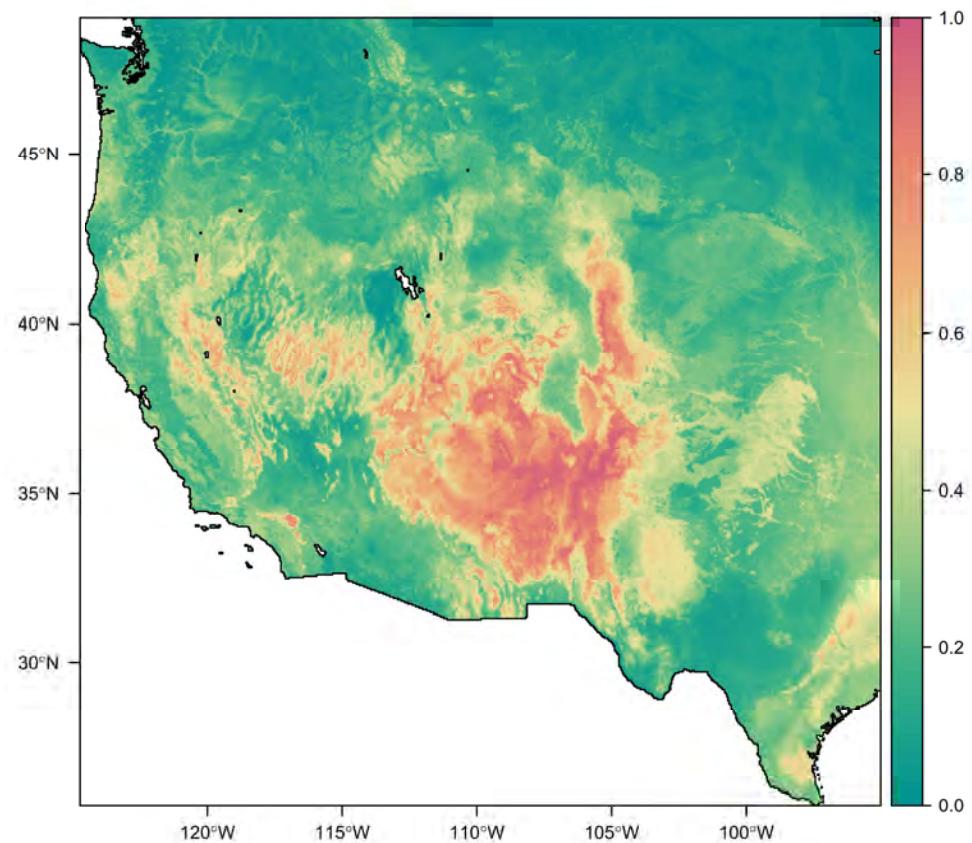
Extended Data Figure 18: Total estimated change in mean temperature, 1950 to present; the region as a whole experienced an average warming of 0.84 °C.



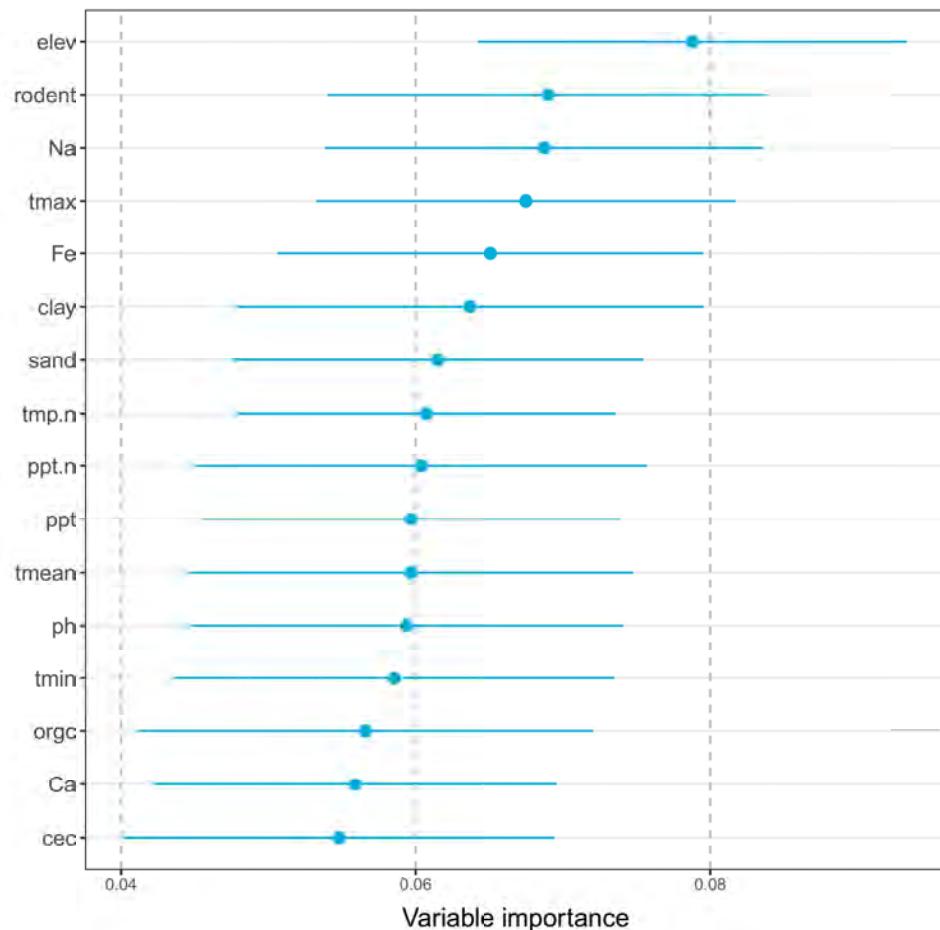
Extended Data Figure 19: Total estimated change in precipitation, 1950 to present; the region as a whole experienced an average increase of 41.7 mm of annual precipitation.



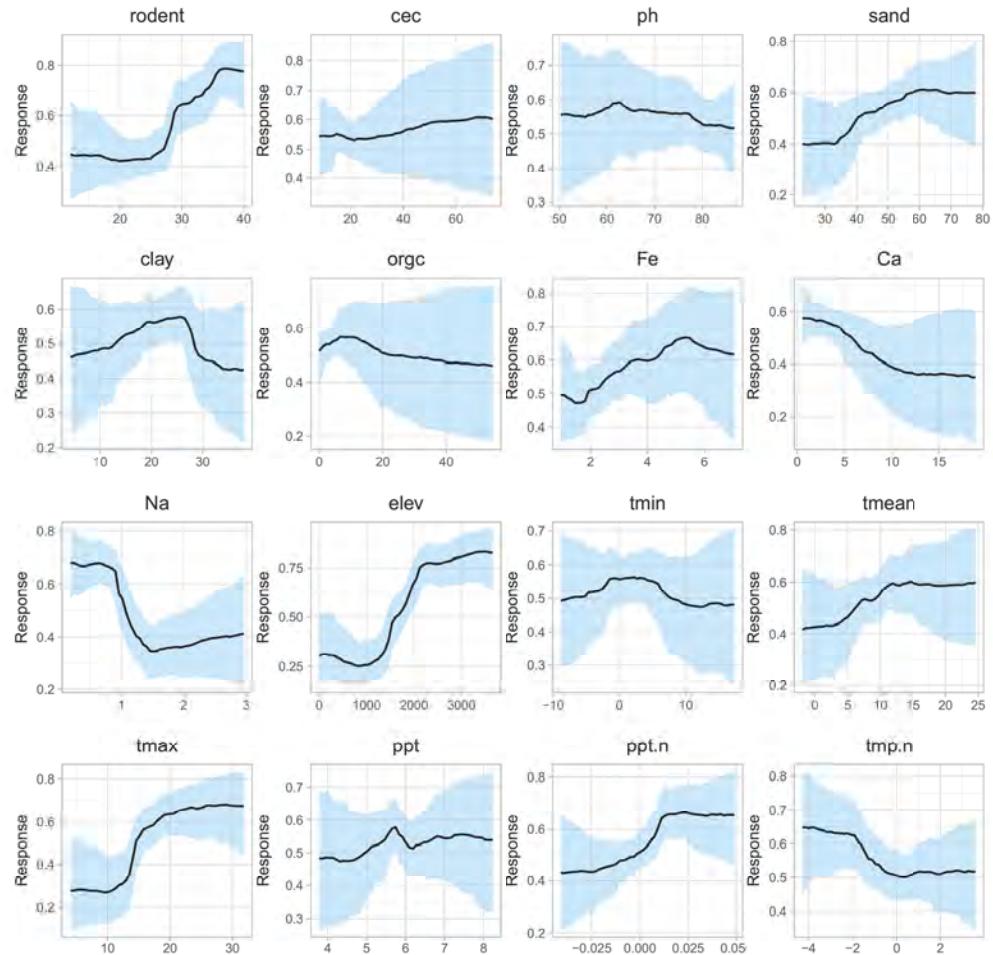
Extended Data Figure 20: Mean suitability for plague across all years (1950-2017), using the alternate human model (no variable set reduction).



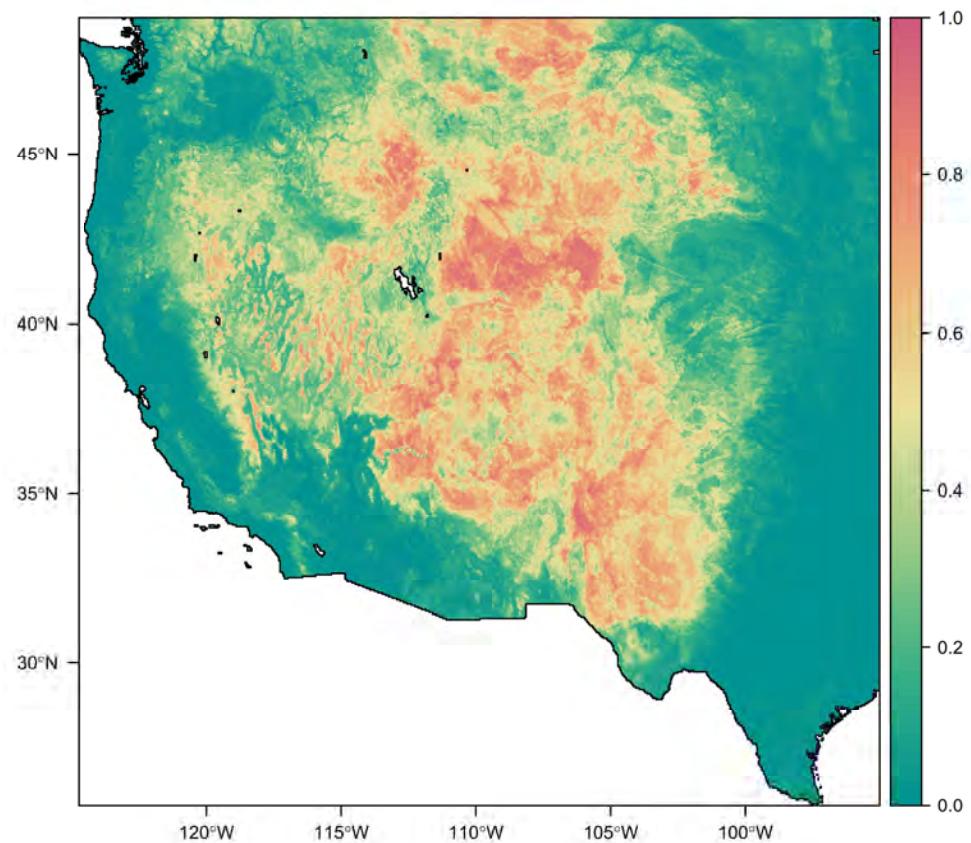
Extended Data Figure 21: The variable importance diagnostic for the alternate human model using the full variables set (no variable set reduction).



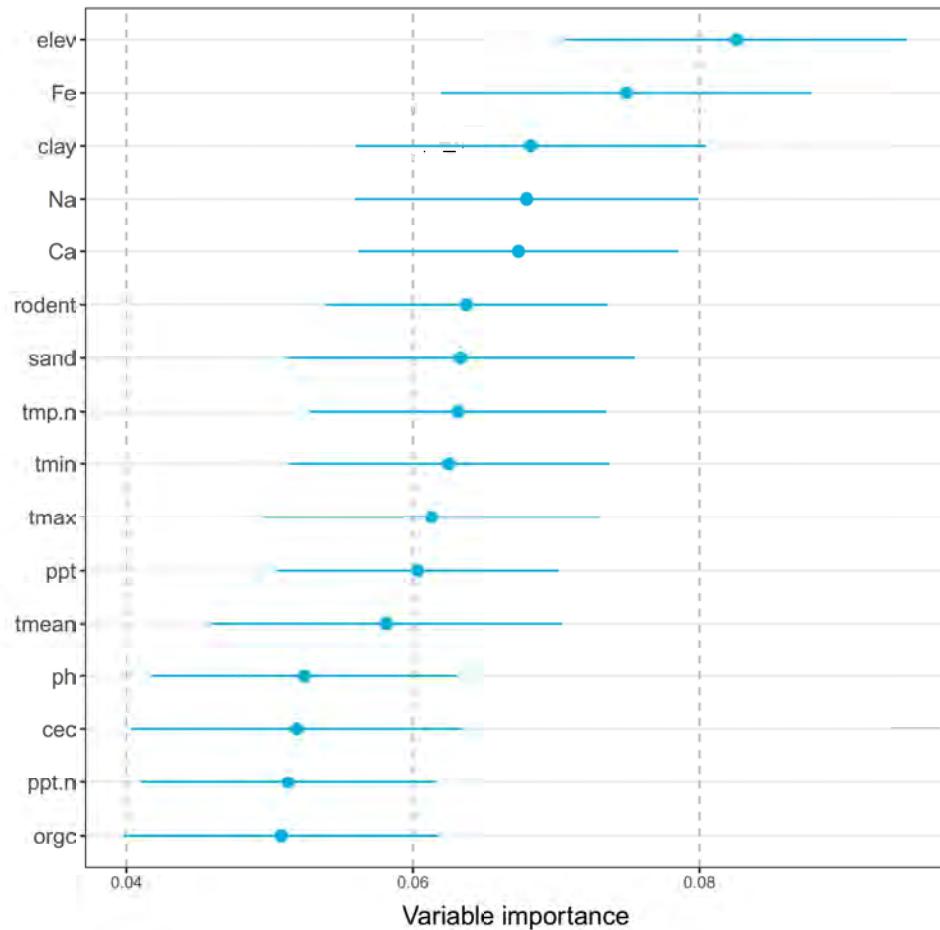
Extended Data Figure 22: Partial dependence plots for the alternate human model using the full variables set (no variable set reduction).



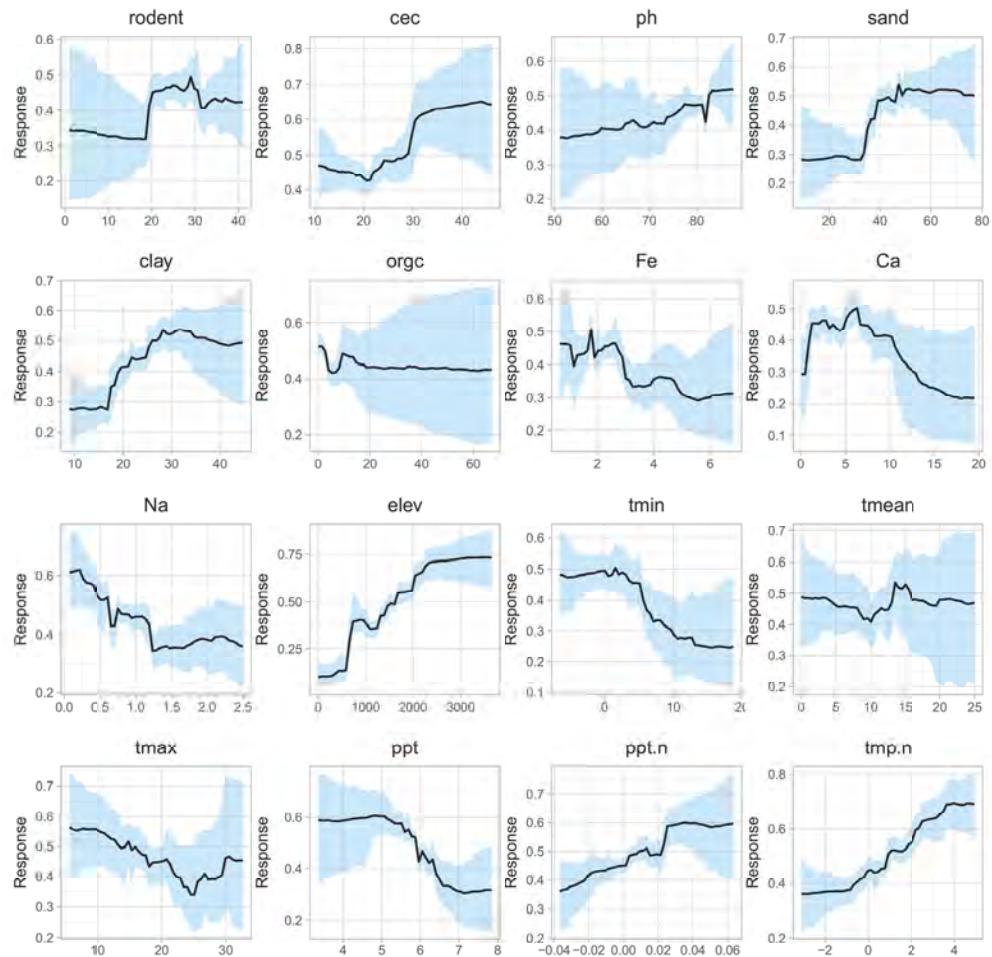
Extended Data Figure 23: Mean suitability for plague across all years (1950-2017), using the first alternate wildlife model (no variable set reduction).



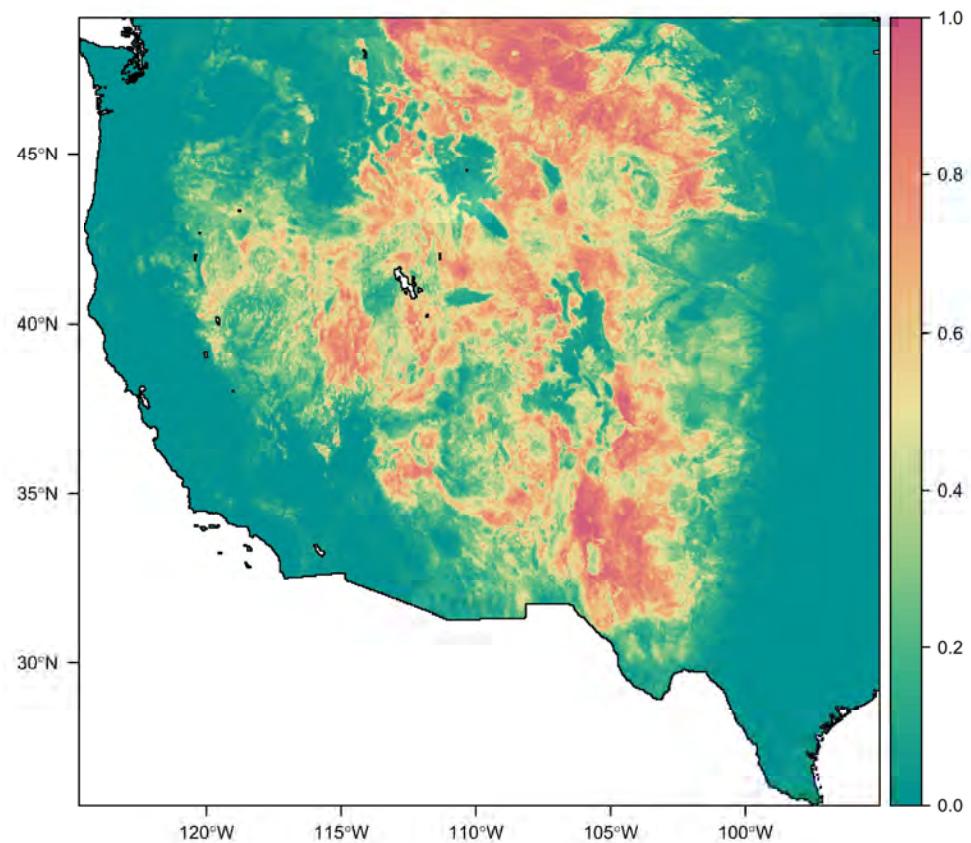
Extended Data Figure 24: The variable importance diagnostic for the first alternate wildlife model using the full variables set (no variable set reduction).



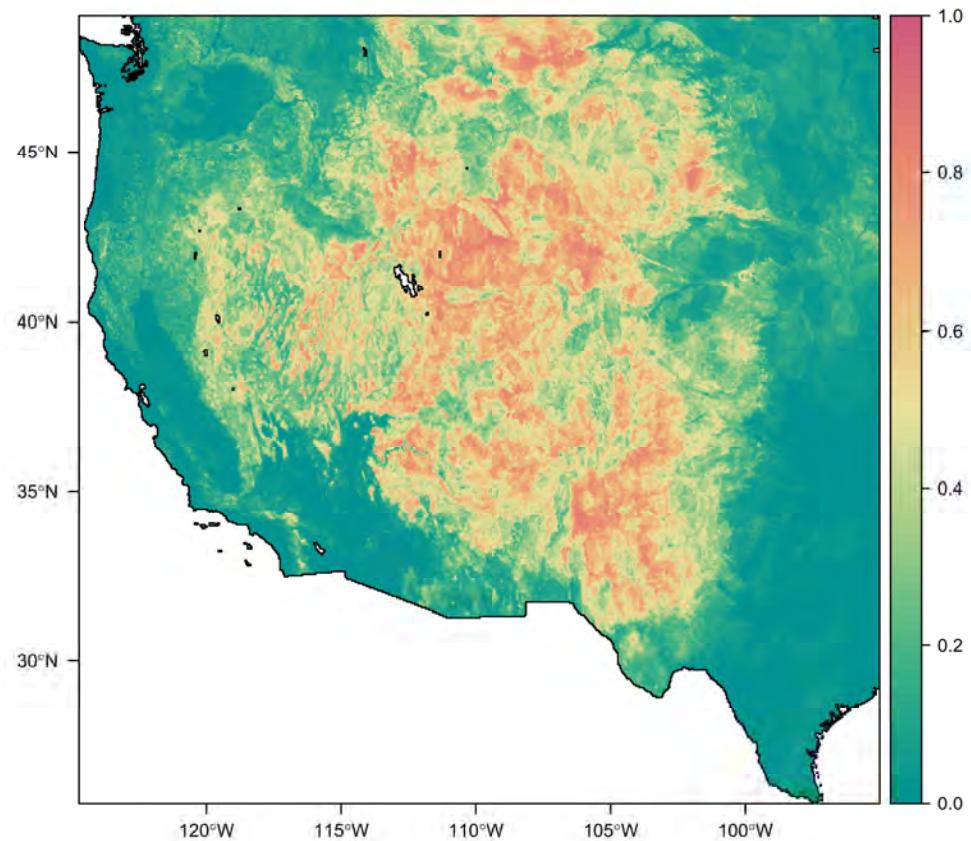
Extended Data Figure 25: Partial dependence plots for the alternate wildlife model using the full variables set (no variable set reduction).



Extended Data Figure 26: Mean suitability for plague across all years (1950-2017), using the second alternate wildlife model (pseudoabsences instead of true absences).



Extended Data Figure 27: Mean suitability for plague across all years (1950-2017), using the second alternate wildlife model (only coyote data).



Extended Data Table 1: Variable name abbreviations used in this study, with full variable names and descriptions.

Variable	Full variable name
rodent	Rodent species richness
cec	Soil cation exchange capacity
ph	Soil pH (acidity)
sand	Soil percent sand content by volume
clay	Soil percent clay content by volume
org	Soil organic carbon content
Fe	Soil iron macronutrient concentration
Ca	Soil calcium macronutrient concentration
Na	Soil sodium macronutrient concentration (≈ proxy for salinity)
elev	Elevation above sea level
tmin	Average minimum annual temperature
tmean	Mean annual temperature
tmax	Average maximum annual temperature
ppt	Mean annual precipitation
ppt.n	Annual normalized precipitation anomaly (relative to long-term average)
tmp.n	Annual normalized temperature anomaly (relative to long-term average)

Extended Data Table 2: Methodologies of previous studies applying ecological niche modeling to map plague (*Yersinia pestis*); this excludes studies focused on mapping individual reservoirs or fleas without any plague data. Abbreviations: PET/AET = potential and actual evapotranspiration; NDVI = normalized difference vegetation index; EVI = enhanced vegetation index; CTI = compound topographic index. [†] Note that this study also bootstrapped county level human cases in the U.S.

Study	Extent	Years	Algorithm	Predictors
15 †	USA	1965-2003?	GARP	Precipitation, temperature (min, mean, max), PET, AET, moisture surplus, moisture deficit, slope, aspect, elevation, CTI
42	Africa	1970-2007	GARP	BIO 1, 2, 5, 6, 12-14, slope, aspect, elevation, CTI, soil pH, soil moisture, soil carbon, PET, AET, humidity, growing degree days, NDVI
54	California, USA	1984-2004	MaxEnt	BIO 5, 7, 15-18
109	“North America”	Unspecified	Maxent	BIO 1, 2, 5, 6, 12-14
59	Western Usambara Mountains, Tanzania	1986-2003	GARP	EVI (mean, dry/rainy period means, standard error, seasonality, heterogeneity), slope, aspect, elevation, CTI
63	Northeast Brazil	1966-2011	GARP	BIO 1, 2, 5, 6, 12-14, NDVI, elevation
110	China	1772-1964	GARP	BIO 1, 2, 3, 7, 12, 14, 15
58	Qinghai-Tibetan Plateau, China	2004-2010	MaxEnt	Elevation, land surface temperature (day & night), NDVI, slope, aspect, land cover
57	Western USA	2000-2015	MaxEnt	BIO 1, 11, 16, 17, distances to grassland, shrubland, cropland, sparse vegetation, bare soil, artificial surface, probability of deer mouse, altitude

678 **References**

- 679 1. Carlson, C. J. *et al.* Climate change will drive novel cross-species viral transmission.
680 *bioRxiv* (2020).
- 681 2. Estrada-Peña, A., Ostfeld, R. S., Peterson, A. T., Poulin, R. & de la Fuente, J. Effects
682 of environmental change on zoonotic disease risk: an ecological primer. *Trends in*
683 *parasitology* **30**, 205–214 (2014).
- 684 3. Caminade, C., McIntyre, K. M. & Jones, A. E. Impact of recent and future climate
685 change on vector-borne diseases. *Annals of the New York Academy of Sciences* **1436**,
686 157 (2019).
- 687 4. Rocklöv, J. & Dubrow, R. Climate change: an enduring challenge for vector-borne
688 disease prevention and control. *Nature Immunology* **21**, 479–483 (2020).
- 689 5. Gibb, R. *et al.* Zoonotic host diversity increases in human-dominated ecosystems.
690 *Nature* **584**, 398–402 (2020).
- 691 6. Loh, E. H. *et al.* Targeting transmission pathways for emerging zoonotic disease
692 surveillance and control. *Vector-Borne and Zoonotic Diseases* **15**, 432–437 (2015).
- 693 7. Carleton, T. A. & Hsiang, S. M. Social and economic impacts of climate. *Science*
694 **353** (2016).
- 695 8. Ebi, K. L., Ogden, N. H., Semenza, J. C. & Woodward, A. Detecting and attributing
696 health burdens to climate change. *Environmental health perspectives* **125**, 085004
697 (2017).
- 698 9. Escobar, L. E. & Craft, M. E. Advances and limitations of disease biogeography using
699 ecological niche modeling. *Frontiers in Microbiology* **7**, 1174 (2016).
- 700 10. Hay, S. I. *et al.* Global mapping of infectious disease. *Philosophical Transactions of*
701 *the Royal Society B: Biological Sciences* **368**, 20120250 (2013).
- 702 11. Blackburn, J. K., McNyset, K. M., Curtis, A. & Hugh-Jones, M. E. Modeling the geo-
703 graphic distribution of bacillus anthracis, the causative agent of anthrax disease, for
704 the contiguous united states using predictive ecologic niche modeling. *The American*
705 *journal of tropical medicine and hygiene* **77**, 1103–1110 (2007).
- 706 12. Trisos, C. H., Merow, C. & Pigot, A. L. The projected timing of abrupt ecological
707 disruption from climate change. *Nature* **580**, 496–501 (2020).
- 708 13. Dobrowski, S. Z. *et al.* Modeling plant ranges over 75 years of climate change in
709 California, USA: temporal transferability and species traits. *Ecological Monographs*
710 **81**, 241–257 (2011).

- 711 14. Kulkarni, M. A. *et al.* 10 years of environmental change on the slopes of mount
712 kilimanjaro and its associated shift in malaria vector distributions. *Frontiers in public*
713 *health* **4**, 281 (2016).
- 714 15. Nakazawa, Y. *et al.* Climate change effects on plague and tularemia in the united
715 states. *Vector-borne and Zoonotic Diseases* **7**, 529–540 (2007).
- 716 16. Araújo, M. B., Pearson, R. G. & Rahbek, C. Equilibrium of species' distributions
717 with climate. *Ecography* **28**, 693–695 (2005).
- 718 17. Gallien, L., Douzet, R., Pratte, S., Zimmermann, N. E. & Thuiller, W. Invasive
719 species distribution models—how violating the equilibrium assumption can create new
720 insights. *Global Ecology and Biogeography* **21**, 1126–1136 (2012).
- 721 18. Peterson, A. T., Martínez-Campos, C., Nakazawa, Y. & Martínez-Meyer, E. Time-
722 specific ecological niche modeling predicts spatial dynamics of vector insects and hu-
723 man dengue cases. *Transactions of the Royal Society of Tropical Medicine and Hygiene*
724 **99**, 647–655 (2005).
- 725 19. Kaul, R. B., Evans, M. V., Murdock, C. C. & Drake, J. M. Spatio-temporal spillover
726 risk of yellow fever in brazil. *Parasites & Vectors* **11**, 488 (2018).
- 727 20. Schmidt, J. P. *et al.* Spatiotemporal fluctuations and triggers of ebola virus spillover.
728 *Emerging Infectious Diseases* **23**, 415 (2017).
- 729 21. Bramanti, B., Dean, K. R., Walløe, L. & Chr. Stenseth, N. The third plague pandemic
730 in europe. *Proceedings of the Royal Society B* **286**, 20182429 (2019).
- 731 22. Xu, L. *et al.* Wet climate and transportation routes accelerate spread of human plague.
732 *Proceedings of the Royal Society B: Biological Sciences* **281**, 20133159 (2014).
- 733 23. Jones, S. D. *et al.* Living with plague: Lessons from the soviet union's antiplague
734 system. *Proceedings of the National Academy of Sciences* **116**, 9155–9163 (2019).
- 735 24. Neerinckx, S., Bertherat, E. & Leirs, H. Human plague occurrences in africa: an
736 overview from 1877 to 2008. *Transactions of the Royal Society of Tropical Medicine
737 and Hygiene* **104**, 97–103 (2010).
- 738 25. Seal, S. Epidemiological studies of plague in india: 1. the present position. *Bulletin
739 of the World Health Organization* **23**, 283 (1960).
- 740 26. Link, V. B. *A history of plague in the United States of America*. 392 (US Government
741 Printing Office, 1955).
- 742 27. Morelli, G. *et al.* Yersinia pestis genome sequencing identifies patterns of global
743 phylogenetic diversity. *Nature genetics* **42**, 1140–1143 (2010).

- 744 28. Schneider, M. C. *et al.* Where does human plague still persist in latin america? *PLoS*
745 *Negl Trop Dis* **8**, e2680 (2014).
- 746 29. Sun, Z. *et al.* Human plague system associated with rodent diversity and other
747 environmental factors. *Royal Society Open Science* **6**, 190216 (2019).
- 748 30. Halliday, F. W. & Rohr, J. R. Measuring the shape of the biodiversity-disease rela-
749 tionship across systems reveals new findings and key gaps. *Nature communications*
750 **10**, 1–10 (2019).
- 751 31. Luis, A. D., Kuenzi, A. J. & Mills, J. N. Species diversity concurrently dilutes and
752 amplifies transmission in a zoonotic host-pathogen system through competing mech-
753 anisms. *Proceedings of the National Academy of Sciences* **115**, 7979–7984 (2018).
- 754 32. Benavides-Montaño, J. A. & Vadyvaloo, V. Yersinia pestis resists predation by acan-
755 thamoeba castellanii and exhibits prolonged intracellular survival. *Applied and envi-
756 ronmental microbiology* **83** (2017).
- 757 33. Boegler, K. A. *et al.* Evaluation of the infectiousness to mice of soil contaminated
758 with yersinia pestis-infected blood. *Vector-Borne and Zoonotic Diseases* **12**, 948–952
759 (2012).
- 760 34. Ayyadurai, S. *et al.* Long-term persistence of virulent yersinia pestis in soil. *Micro-
761 biology* **154**, 2865–2871 (2008).
- 762 35. Baltazard, M., Karimi, Y., Eftekhari, M., Chamsa, M. & Mollaret, H. Interepidzootic
763 conservation of plague in its inveterate foci: working hypotheses. Tech. Rep., ARMY
764 BIOLOGICAL LABS FREDERICK MD (1964).
- 765 36. Eisen, R. J. *et al.* Persistence of yersinia pestis in soil under natural conditions.
766 *Emerging infectious diseases* **14**, 941 (2008).
- 767 37. Karimi, Y. Natural preservation of plague in soil. *Bulletin de la Societe de pathologie
768 exotique et de ses filiales* **56**, 1183–1186 (1963).
- 769 38. Carlson, C. J. *et al.* The global distribution of bacillus anthracis and associated
770 anthrax risk to humans, livestock and wildlife. *Nature Microbiology* **1** (2019).
- 771 39. Limmathurotsakul, D. *et al.* Predicted global distribution of burkholderia pseudoma-
772 llei and burden of melioidosis. *Nature Microbiology* **1**, 15008 (2016).
- 773 40. Malek, M. A. *et al.* Yersinia pestis halotolerance illuminates plague reservoirs. *Sci-
774 entific reports* **7**, 1–10 (2017).
- 775 41. Barbieri, R., Texier, G., Keller, C. & Drancourt, M. Soil salinity and aridity specify
776 plague foci in the united states of america. *Scientific reports* **10**, 1–9 (2020).

- 777 42. Neerinckx, S. B., Peterson, A. T., Gulinck, H., Deckers, J. & Leirs, H. Geographic
778 distribution and ecological niche of plague in sub-saharan africa. *International Journal*
779 *of Health Geographics* **7**, 54 (2008).
- 780 43. Bacot, A. Lxix. a study of the bionomics of the common rat fleas and other species
781 associated with human habitations, with special reference to the influence of temper-
782 ature and humidity at various periods of the life history of the insect. *The Journal of*
783 *hygiene* **13**, 447 (1914).
- 784 44. Eisen, R. J., Eisen, L. & Gage, K. L. Studies of vector competency and efficiency of
785 north american fleas for yersinia pestis: state of the field and future research needs.
786 *Journal of medical entomology* **46**, 737–744 (2009).
- 787 45. Bland, D. M., Jarrett, C. O., Bosio, C. F. & Hinnebusch, B. J. Infectious blood
788 source alters early foregut infection and regurgitative transmission of yersinia pestis
789 by rodent fleas. *PLoS pathogens* **14**, e1006859 (2018).
- 790 46. Cui, Y. *et al.* Evolutionary selection of biofilm-mediated extended phenotypes in
791 yersinia pestis in response to a fluctuating environment. *Nature communications* **11**,
792 1–8 (2020).
- 793 47. Parmenter, R. R., Yadav, E. P., Parmenter, C. A., Ettestad, P. & Gage, K. L. Inci-
794 dence of plague associated with increased winter-spring precipitation in new mexico.
795 *The American journal of tropical medicine and hygiene* **61**, 814–821 (1999).
- 796 48. Xu, L. *et al.* The trophic responses of two different rodent–vector–plague systems to
797 climate change. *Proceedings of the Royal Society B: Biological Sciences* **282**, 20141846
798 (2015).
- 799 49. Kausrud, K. *et al.* Climatically driven synchrony of gerbil populations allows large-
800 scale plague outbreaks. *Proceedings of the Royal Society B: Biological Sciences* **274**,
801 1963–1969 (2007).
- 802 50. Adjemian, J. Z., Foley, P., Gage, K. L. & Foley, J. E. Initiation and spread of traveling
803 waves of plague, yersinia pestis, in the western united states. *The American Journal*
804 *of Tropical Medicine and Hygiene* **76**, 365–375 (2007).
- 805 51. Ben Ari, T. *et al.* Human plague in the usa: the importance of regional and local
806 climate. *Biology Letters* **4**, 737–740 (2008).
- 807 52. Bevins, S. N., Baroch, J. A., Nolte, D. L., Zhang, M. & He, H. Yersinia pestis:
808 examining wildlife plague surveillance in china and the usa. *Integrative Zoology* **7**,
809 99–109 (2012).

- 810 53. Carlson, C. J. embarcadero: Species distribution modelling with bayesian additive
811 regression trees in r. *BioRxiv* 774604 (2019).
- 812 54. Holt, A. C., Salkeld, D. J., Fritz, C. L., Tucker, J. R. & Gong, P. Spatial analysis
813 of plague in california: niche modeling predictions of the current distribution and
814 potential response to climate change. *International Journal of Health Geographics* **8**,
815 38 (2009).
- 816 55. Brown, H. E. *et al.* Climatic predictors of the intra-and inter-annual distributions of
817 plague cases in new mexico based on 29 years of animal-based surveillance data. *The*
818 *American journal of tropical medicine and hygiene* **82**, 95–102 (2010).
- 819 56. Eisen, R. J. *et al.* Human plague in the southwestern united states, 1957–2004: spatial
820 models of elevated risk of human exposure to *yersinia pestis*. *Journal of Medical*
821 *Entomology* **44**, 530–537 (2007).
- 822 57. Walsh, M. & Haseeb, M. Modeling the ecologic niche of plague in sylvan and domestic
823 animal hosts to delineate sources of human exposure in the western united states.
824 *PeerJ* **3**, e1493 (2015).
- 825 58. Qian, Q. *et al.* Mapping risk of plague in qinghai-tibetan plateau, china. *BMC*
826 *Infectious Diseases* **14**, 382 (2014).
- 827 59. Neerinckx, S. *et al.* Predicting potential risk areas of human plague for the western
828 usambara mountains, lushoto district, tanzania. *The American Journal of Tropical*
829 *Medicine and Hygiene* **82**, 492–500 (2010).
- 830 60. Moore, S. M. *et al.* Seasonal fluctuations of small mammal and flea communities in a
831 ugandan plague focus: evidence to implicate *arvicanthis niloticus* and *crocidura* spp.
832 as key hosts in *yersinia pestis* transmission. *Parasites & vectors* **8**, 1–15 (2015).
- 833 61. Eisen, R. J. *et al.* Flea diversity as an element for persistence of plague bacteria in
834 an east african plague focus. *PLoS One* **7**, e35598 (2012).
- 835 62. Abedi, A. A. *et al.* Ecologic features of plague outbreak areas, democratic republic
836 of the congo, 2004–2014. *Emerging infectious diseases* **24**, 210 (2018).
- 837 63. Giles, J., Peterson, A. T. & Almeida, A. Ecology and geography of plague transmission
838 areas in northeastern brazil. *PLoS Neglected Tropical Diseases* **5**, e925 (2011).
- 839 64. Giorgi, E. *et al.* Modeling of spatio-temporal variation in plague incidence in mada-
840 gascar from 1980 to 2007. *Spatial and spatio-temporal epidemiology* **19**, 125–135
841 (2016).

- 842 65. Kutyrev, V. V. & Popova, A. Y. (eds.) *The Inventory of Epidemic and Epizootic*
843 *Manifestations of Plague on the Territory of the Former Soviet Union from 1876 to*
844 *2016*. Saratov State University. (Saratov State University, 2016).
- 845 66. Rohr, J. R. *et al.* Towards common ground in the biodiversity–disease debate. *Nature*
846 *ecology & evolution* **4**, 24–33 (2020).
- 847 67. Johnson, P. T., Preston, D. L., Hoverman, J. T. & Richgels, K. L. Biodiversity
848 decreases disease through predictable changes in host community competence. *Nature*
849 **494**, 230–233 (2013).
- 850 68. Miller, E. & Huppert, A. The effects of host diversity on vector-borne disease: the
851 conditions under which diversity will amplify or dilute the disease risk. *PLoS One* **8**,
852 e80279 (2013).
- 853 69. Schmidt, K. A. & Ostfeld, R. S. Biodiversity and the dilution effect in disease ecology.
854 *Ecology* **82**, 609–619 (2001).
- 855 70. Mahmoudi, A. *et al.* Plague reservoir species throughout the world. *Integrative*
856 *Zoology* (2020).
- 857 71. Buck, J. & Perkins, S. Study scale determines whether wildlife loss protects against or
858 promotes tick-borne disease. *Proceedings of the Royal Society B: Biological Sciences*
859 **285**, 20180218 (2018).
- 860 72. Mordecai, E. A. *et al.* Thermal biology of mosquito-borne disease. *Ecology letters* **22**,
861 1690–1708 (2019).
- 862 73. Schotthoefer, A. M. *et al.* Effects of temperature on early-phase transmission of
863 *yersinia pestis* by the flea, *xenopsylla cheopis*. *Journal of medical entomology* **48**,
864 411–417 (2011).
- 865 74. Williams, S. K. *et al.* Effects of low-temperature flea maintenance on the transmission
866 of *yersinia pestis* by *oropsylla montana*. *Vector-Borne and Zoonotic Diseases* **13**, 468–
867 478 (2013).
- 868 75. Xu, L. *et al.* Historical and genomic data reveal the influencing factors on global
869 transmission velocity of plague during the third pandemic. *Proceedings of the National*
870 *Academy of Sciences* **116**, 11833–11838 (2019).
- 871 76. Stenseth, N. C. *et al.* Plague dynamics are driven by climate variation. *Proceedings*
872 *of the National Academy of Sciences* **103**, 13110–13115 (2006).
- 873 77. Frigessi, A. *et al.* Bayesian population dynamics of interacting species: great gerbils
874 and fleas in kazakhstan. *Biometrics* **61**, 230–238 (2005).

- 875 78. Samia, N. I. *et al.* Dynamics of the plague–wildlife–human system in central asia are
876 controlled by two epidemiological thresholds. *Proceedings of the National Academy of*
877 *Sciences* **108**, 14527–14532 (2011).
- 878 79. Reijniers, J., Begon, M., Ageyev, V. S. & Leirs, H. Plague epizootic cycles in central
879 asia. *Biology letters* **10**, 20140302 (2014).
- 880 80. Schmid, B. V. *et al.* Climate-driven introduction of the black death and successive
881 plague reintroductions into europe. *Proceedings of the National Academy of Sciences*
882 **112**, 3020–3025 (2015).
- 883 81. Markman, D. W. *et al.* Yersinia pestis survival and replication in potential ameba
884 reservoir. *Emerging infectious diseases* **24**, 294 (2018).
- 885 82. Downs, C. J., Pinshow, B., Khokhlova, I. S. & Krasnov, B. R. Flea fitness is reduced
886 by high fractional concentrations of co₂ that simulate levels found in their hosts'
887 burrows. *Journal of Experimental Biology* **218**, 3596–3603 (2015).
- 888 83. Brickner-Braun, I. *et al.* Ventilation of multi-entranced rodent burrows by boundary
889 layer eddies. *Journal of Experimental Biology* **217**, 4141–4148 (2014).
- 890 84. Warth, J. F. G., Biesdorf, S. M. & de Souza, C. Yersinia pseudotuberculosis o iii causes
891 diarrhea in brazilian cattle. In *Advances in Yersinia Research*, 107–110 (Springer,
892 2012).
- 893 85. Ell, S. R. Iron in two seventeenth-century plague epidemics. *The Journal of interdis-*
894 *ciplinary history* **15**, 445–457 (1985).
- 895 86. Hooker-Romero, D. *et al.* Iron availability and oxygen tension regulate the yersinia
896 ysc type iii secretion system to enable disseminated infection. *PLoS pathogens* **15**,
897 e1008001 (2019).
- 898 87. Quenee, L. E. *et al.* Hereditary hemochromatosis restores the virulence of plague
899 vaccine strains. *The Journal of infectious diseases* **206**, 1050–1058 (2012).
- 900 88. Fowler, J. M., Wulff, C. R., Straley, S. C. & Brubaker, R. R. Growth of calcium-
901 blind mutants of yersinia pestis at 37 °c in permissive ca²⁺-deficient environments.
902 *Microbiology* **155**, 2509 (2009).
- 903 89. Schneewind, O. Classic spotlight: Studies on the low-calcium response of yersinia
904 pestis reveal the secrets of plague pathogenesis. *Journal of bacteriology* **198**, 2018
905 (2016).
- 906 90. Carlson, C. J. *et al.* Spores and soil from six sides: interdisciplinarity and the envi-
907 ronmental biology of anthrax (*bacillus anthracis*). *Biological Reviews* **93**, 1813–1831
908 (2018).

- 909 91. Liccioli, S. *et al.* Enzootic maintenance of sylvatic plague in canada's threatened
910 black-tailed prairie dog ecosystem. *Ecosphere* **11** (2020).
- 911 92. Kraemer, M. U. *et al.* Progress and challenges in infectious disease cartography.
912 *Trends in Parasitology* **32**, 19–29 (2016).
- 913 93. Stenseth, N. C. *et al.* Plague: past, present, and future. *PLoS Medicine* **5**, e3 (2008).
- 914 94. Brown, H. E. *et al.* Annual seroprevalence of yersinia pestis in coyotes as predictors
915 of interannual variation in reports of human plague cases in arizona, united states.
916 *Vector-Borne and Zoonotic Diseases* **11**, 1439–1446 (2011).
- 917 95. Gage, K., Montenieri, J. & Thomas, R. The role of predators in the ecology, epidemi-
918 ology, and surveillance of plague in the united states. In *Vertebrate Pest Conference*
919 (*USA*) (1994).
- 920 96. Chu, C. M. *Laboratory Manual Of Plague Diagnostic Tests*. Center for Disease Con-
921 trol and Prevention. (U.S. Centers for Disease Control and Prevention, 2000).
- 922 97. Chandler, J. C. *et al.* A bead-based flow cytometric assay for monitoring yersinia
923 pestis exposure in wildlife. *Journal of clinical microbiology* **56** (2018).
- 924 98. Di Luzio, M., Johnson, G. L., Daly, C., Eischeid, J. K. & Arnold, J. G. Constructing
925 retrospective gridded daily precipitation and temperature datasets for the conterminous
926 united states. *Journal of Applied Meteorology and Climatology* **47**, 475–497
927 (2008).
- 928 99. Hengl, T. *et al.* Soilgrids250m: Global gridded soil information based on machine
929 learning. *PloS one* **12**, e0169748–e0169748 (2017).
- 930 100. Smith, D. B., Cannon, W. F. & Woodruff, L. G. A national-scale geochemical and
931 mineralogical survey of soils of the conterminous united states. *Applied Geochemistry*
932 **26**, S250–S255 (2011).
- 933 101. Smith, D. B. *et al.* Geochemical and mineralogical data for soils of the conterminous
934 united states. Tech. Rep., US Geological Survey (2013).
- 935 102. Norberg, A. *et al.* A comprehensive evaluation of predictive performance of 33 species
936 distribution models at species and community levels. *Ecological Monographs* e01370
937 (2019).
- 938 103. Bhatt, S. *et al.* The global distribution and burden of dengue. *Nature* **496**, 504–507
939 (2013).

- 940 104. Hieronimo, P. *et al.* Integrating land cover and terrain characteristics to explain plague
941 risks in western usambara mountains, tanzania: a geospatial approach. *Tanzania*
942 *journal of health research* **16** (2014).
- 943 105. Pigott, D. M. *et al.* Mapping the zoonotic niche of ebola virus disease in africa. *Elife*
944 **3**, e04395 (2014).
- 945 106. Richards, R. L. *et al.* Identifying correlates of guinea worm (dracunculus medinensis)
946 infection in domestic dog populations. *PLOS Neglected Tropical Diseases* **14**, e0008620
947 (2020).
- 948 107. Shearer, F. M. *et al.* Existing and potential infection risk zones of yellow fever world-
949 wide: a modelling analysis. *The Lancet Global Health* **6**, e270–e278 (2018).
- 950 108. Chipman, H. A., George, E. I., McCulloch, R. E. *et al.* Bart: Bayesian additive
951 regression trees. *The Annals of Applied Statistics* **4**, 266–298 (2010).
- 952 109. Maher, S. P., Ellis, C., Gage, K. L., Enscore, R. E. & Peterson, A. T. Range-
953 wide determinants of plague distribution in north america. *The American Journal of*
954 *Tropical Medicine and Hygiene* **83**, 736–742 (2010).
- 955 110. Ben-Ari, T. *et al.* Identification of chinese plague foci from long-term epidemiological
956 data. *Proceedings of the National Academy of Sciences* **109**, 8196–8201 (2012).