ORIGINAL PAPER

Predicted threats to a native squirrel from two invading species based on citizen science data

Theodore T. Tran · Benjamin E. Carter [ORCID] ·
Jessica A. Castillo Vardaro [ORCID]

Abstract Invasive species are a major threat to biodiversity worldwide and tree squirrels (Sciuridae) are among the most successful mammalian invaders. Two tree squirrels native to the eastern United States, the eastern grey squirrel (*Sciurus carolinensis*) and eastern fox squirrel (*Sciurus niger*), have been repeatedly introduced into the western United States over the past century. In that time, the non-native species have expanded their ranges extensively and have negatively impacted the only native species of large tree squirrel in the western United States, the western grey squirrel (*Sciurus griseus*). While numerous studies have documented the impacts of these invasions on *S. griseus*, few have assessed the potential for future expansion and none have evaluated all three squirrel species simultaneously. In this study, we address: whether the predicted distributions derived from observations in the western states differ from those derived from the native eastern distribution; to what extent do the predicted distributions overlap with the native *S. griseus*; what role does urbanization play in invasion potential of the introduced squirrels. We use citizen science data from the iNaturalist database to model the species distributions in the western United States for all three *Sciurus* species. We generated models based on occurrences in the squirrels' native ranges, compared to models generated from their introduced ranges, and evaluated current and predicted conflict zones. We determined that the greatest potential conflict with the native squirrel is currently in areas adjacent to regions of high human footprint. As human development expands, the invasive squirrels are likely to expand into previously inaccessible areas, increasing conflict with and potentially displacing the native western grey squirrel.

Keywords Invasive species · Species distribution model · *Sciurus niger* · *Sciurus carolinensis* · *Sciurus griseus*

T. T. Tran · B. E. Carter · J. A. Castillo Vardaro (✉)
Department of Biological Sciences, San José State University, San Jose, CA, USA
e-mail: Jessica.Castillo-Vardaro@sjsu.edu

Introduction

Introduction of non-native species is among the leading threats to biodiversity worldwide (Sala et al. 2000). Although interacting effects of invasions with climate change and habitat loss are predicted to be particularly dire (Hellmann et al. 2008), new tools and datasets are emerging that facilitate our ability to understand the complex array of factors driving invasions. The citizen science platform iNaturalist is a particularly good example of this progress, especially in the context of monitoring invasions. Citizen science data tends to be most dense in exactly the urban/

wildland interfaces where invasion fronts are most active and, for charismatic organisms like vertebrates and showy flowering plants, data can be abundant.

One group with particularly high promise for studying invasions using citizen science data is the squirrel family (Sciuridae). Squirrels are photogenic, charismatic, diurnal, and tend not to avoid humans. They also can be particularly tenacious invaders, with up to 80% of squirrel introductions worldwide leading to the successful establishment of a viable breeding population (Bertolino 2009). Sciurids have a near cosmopolitan distribution, with many species having been purposefully introduced outside their native range and subsequently implicated in declines of native squirrel species (Bertolino 2009). As with most invasions, the success of squirrel introductions is linked to vegetation type, extent of human development, and food availability (Williamson 1983; Parker and Nilon 2008). Furthermore, ranges of introduced squirrels are predicted to shift over the next 50 years as a result of climate change, with some species greatly expanding their ranges (Di Febbraro et al. 2016). Even small-scale introduction events hold major risks of a viable population establishment. For example, introduction of the North American eastern grey squirrel (*S. carolinensis*) to Italy occurred in 1948 with two males and two females (Bertolino et al. 2000; Martinoli et al. 2010). In 2005 the population had grown to occupy an area of 450 km^2 (Bertolino and Genovesi 2005) and continues to expand (Bertolino et al. 2000), with the expansion apparently linked to local extirpations of populations of the native red squirrel (*S. vulgaris*).

Sciurus carolinensis has also been introduced, along with another eastern North American species (the eastern fox squirrel, *S. niger*), to western North America, where both species have heavily impacted the native western grey squirrel, *S. griseus* (Carraway and Verts 1994; Cooper and Muchlinski 2015; Jessen et al. 2018). *Sciurus niger* was originally introduced into Golden Gate Park (San Francisco, CA), in 1890, but subsequently disappeared (Byrne 1979). Separate introductions occurred in Fresno in 1901, the Los Angeles metropolitan area in 1904, San Diego in 1920, the UC Berkeley Campus in 1926, and Bakersfield in 1985 (Byrne 1979; Claytor et al. 2015). *Sciurus carolinensis* is believed to have been first introduced in the western US on the campus of Stanford University, California in 1939, and to San Francisco's Golden Gate Park in 1943 (Byrne 1979). It currently occurs throughout most of the San Francisco Bay Area and in scattered populations near Sacramento and Stockton in California's Central Valley (Fig. 1). Since their introductions, both species became the dominant arboreal squirrel species in urban areas and have caused the local extirpation of *S. griseus*. For example, *S. niger* has replaced *S. griseus* in multiple lowland oak and evergreen woodland sites throughout Southern California that are directly adjacent to or have become urbanized (Cooper and Muchlinski 2015). Likewise, *S. carolinensis* appears to be displacing *S. griseus* in the Santa Cruz Mountains between Santa Cruz and the densely populated San Francisco Bay Area (Jessen et al. 2018).

Within their native ranges, both *S. carolinesis* and *S. niger* have less specific feeding ecology and habitat associations in comparison to *S. griseus* (Carraway and Verts 1994; Koprowski 1994b, a). *Sciurus carolinesis* and *S. niger* are widespread throughout much of their native ranges in the eastern United States, and are found in a diversity of dense to open evergreen and deciduous forest habitats (Koprowski 1994b, a). Their native ranges overlap considerably (Fig. 1), but *S. carolinensis* typically occupies areas with dense understory while *S. niger* tends to occur in more open habitats (Koprowski 1994b, a). *Sciurus carolinensis* is restricted to more contiguous habitat in its native range while *S. niger* can occur in more fragmented areas (Koprowski 1994b, a). The two species can be syntopic, co-occurring with little conflict, however, there is evidence that one may displace the other under some environmental conditions (Koprowski 1994b). In the western US, *S. griseus* occurs in woodland and forested areas in California, Oregon, Nevada, and Washington (Carraway and Verts 1994).

Both of the introduced species have behavioral characteristics that may promote their success relative to *S. griseus*, in at least some habitats, in western North America. There is a significant overlap in diet between *S. griseus* and *S. niger* in western North America, but *S. niger* consumes a wider range of foods, including many non-native and native plants, while *S. griseus* primarily consumes acorns and pine nuts from native tree species (Ortiz and Muchlinski 2015). This more flexible diet may explain why *S. niger* has been able to outcompete *S. griseus* in disturbed areas and why its survival rates in urban areas are higher than those in rural areas (McCleery et al. 2008). Further, a study of the two species in urban
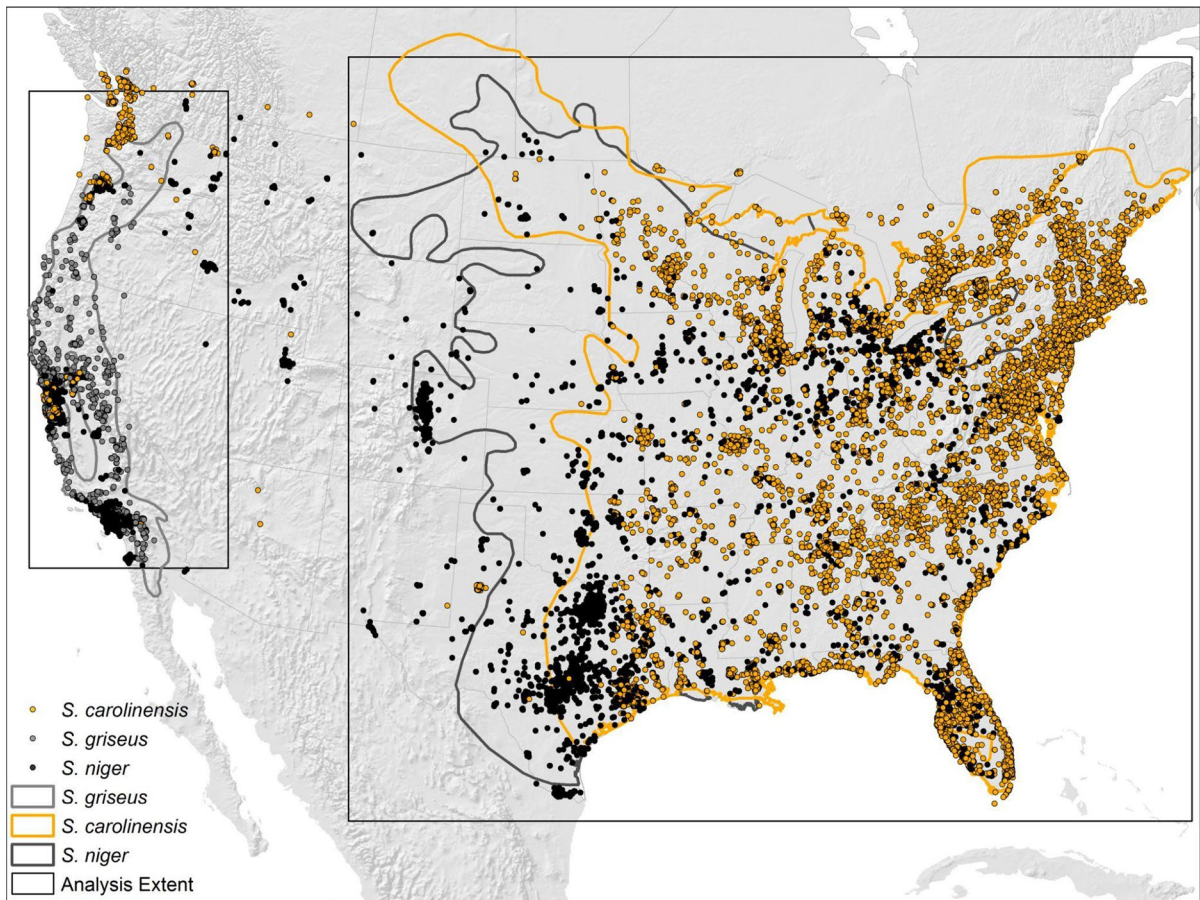
Fig. 1 Map of observations, rarefied to no more than one per 200 m×200 m pixel, used in Maxent analyses. Eastern and western analysis extents are shown in solid black rectangles and the native range for the two eastern species are shown in colored polygons (species range data from The IUCN Red List of Threatened Species. Version 2020–3. http://www.iucnredlist.org)

and suburban parks and natural areas found significant overlap in space and time, which could further contribute to declines in *S. griseus* due to increased competition (Ortiz 2022).

Sciurus carolinensis has not been thoroughly studied ecologically in western North America, but *S. carolinensis'* impacts on native squirrels in Europe are better understood. There, it has a larger foraging range than the native red squirrel (*Sciurus vulgaris*, Wauters et al. 2002), and also lacks anti-predator behaviors demonstrated by *S. vulgaris* (Twining et al. 2020). Documentation of the extent to which *S. niger* and *S. carolinensis* have invaded the western US, and the extent to which they are displacing the native *S. griseus*, remain largely unknown. Here, we take a modeling approach and leverage a large dataset of

citizen science observations to ask the following four questions: (1) Based on their native distributions in eastern North America, are *S. niger* and *S. carolinensis* expected to continue expanding their ranges in the west? (2) Based on their current introduced ranges in the west, are the two species expected to continue expanding their distributions? (3) Are regions of projected expansion of the introduced species coincident with the current distribution of the native *S. griseus* (i.e., should we anticipate further competitive displacement of *S. griseus*, and if so, where)? (4) Do habitat data estimating human impact on the environment (in this case, impervious surfaces and developed areas) significantly improve distribution models over models employing only climate and vegetation type?

Materials and methods

Data collection

We obtained observation data from iNaturalist using the search term "*Sciurus*" on the continental United States collected from 7 March 2008 to 27 March 2020. The observations were filtered to only include sightings deemed "Research Grade" as per iNaturalist guidelines: has a definitive date, is georeferenced, includes at least one clear photo for visual evidence and a ratio 2:3 of identifiers agree on a taxon. These observations were downloaded directly from iNaturalist.org, imported into ArcMap 10.8.1 software (www.esri.com), and overlaid on a land cover raster layer (see Spatial Data, below) to remove any records that fell within water. We retained all records occurring within terrestrial habitat types.

We used two sources of spatial data as our environmental variables for distribution modeling: WorldClim (https://www.worldclim.org) and the National Land Cover Dataset (NLCD) (Dewitz 2019). We used four bioclimatic variables from WorldClim relating to seasonal temperature and precipitation from the years 1970–2000: annual mean temperature (BIO1), maximum temperature of the warmest month (BIO5), minimum temperature of the coldest month (BIO6), and mean annual precipitation (BIO12) at 30-s resolution (~1 km). While the time period for the climate data is prior to the iNaturalist observations, species distributions reflect both current and past conditions and the 30-year normals reflect the averages of the recent past. We included habitat-related data from the NLCD for the year 2016: land cover type, tree canopy cover (percent cover), and the extent of impervious (human-made) surfaces (percentage of developed surface) at 30 m resolution. Land cover type includes 16 classifications within the following categories: water,

developed, barren, forest, shrubland, herbaceous, planted/cultivated, and wetlands (Table S5). All spatial data were projected to WGS84 Albers Equal Area Conical, resampled to a resolution of 200 m × 200 m, and converted to ASCII files for analysis in Maxent. To prevent clustered observations from creating biases in the Maxent model (Boria et al. 2014), we spatially rarefied our squirrel observations to ensure only a single observation per species per pixel, using the SDM toolbox spatial rarefaction function (Brown 2014).

Species distribution model (SDM)

We ran four sets of models on data from the native ranges of three species (*S. carolinensis* and *S niger* in the east, *S. griseus* in the west) and separately from the introduced ranges (*S. niger* and *S. carolinensis* in the west) using Maxent version 3.4.4 (Phillips et al. 2017) for a total of 20 separate models (Table 1). The *Full* model included all four BioClim and three NLCD variables. The *BioClim only* model included only the four BioClim variables. The *BioClim + Land cover* model excluded canopy cover and imperviousness, while the *BioClim, No Land cover* model included canopy cover and imperviousness but excluded land cover. We used the program's default settings and set aside 75% of the observations for training and the remaining 25% for testing the model (Phillips and Dudík 2008). We ran each of the four sets of models for the two eastern species in the eastern extent and projected the predicted distribution probability to the western extent (Fig. 1). Likewise, we ran each set of models for all three squirrel species in the western extent. We evaluated the performance of the models by comparing the area under the curve (AUC) within each species and geographic extent (Merow et al. 2013; Pearson 2010). We then

Table 1 Number of observations for each species by geographic extent, with total observations and observations included in MaxEnt analysis in parentheses. Model support (area under curve, AUC) for each set of variables and geographic extent

Extent	Species	Observations	Full	BioClim only	Bio-Clim + Land cover	BioClim, No land cover
Eastern	*S. carolinensis*	42,209 (16,361)	**0.681**	0.655	0.679	0.680
	S. niger	16,231 (7495)	**0.770**	0.765	0.766	0.766
Western	*S. carolinensis*	3350 (1173)	0.948	0.924	**0.949**	0.948
	S. griseus	2904 (1462)	**0.901**	0.887	0.896	**0.901**
	S. niger	10,790 (3677)	**0.871**	0.860	0.869	0.869

Best supported model is in bold

assessed which of the environmental variables contributed most to the predicted species distribution by comparing the percent contribution and permutation importance for the variables within the best supported model.

To visually characterize where current and predicted conflict among the native and introduced species may occur, we overlaid iNaturalist observations of the introduced species on the predicted distributions derived from the western extent. We created a 1 km radius buffer around each observation, approximately equal to the home range of five to 15 squirrels (Koprowski 1994b, a). We classified the resulting pixels as current conflict (introduced species have been observed and *S. griseus* is predicted), predicted conflict (introduced species and *S. griseus* are both predicted, but introduced has not been observed), invaded without conflict (introduced species has been observed, but *S. griseus* is not predicted), and potential invasion without conflict (introduced species is predicted but *S. griseus* is not).

Post SDM analysis

In order to quantify the degree of predicted overlap among species, as well as the environmental variables contributing most to the predicted overlap, we converted the predicted probability distribution maps to presence absence and conducted linear discriminant function analyses using the *MASS* package (Venables and Ripley 2002) in R (R Core Team) following Rödder and Engler (2011). We transformed the Maxent predicted probability rasters into a matrix with columns for predicted probability of presence for all three species, as well as the input values for each of the environmental variables. We set a threshold of 0.60 to consider a species present for our categorical presence–absence variables, then removed all raster cells that had been projected to be less than 0.60 and thus considered an absence. Previous studies have used lower thresholds, such as the 10th percentile training presence logistic threshold, to predict squirrel distributions in California (Creley et al. 2019). This lower threshold corresponds to values ranging from 0.26 to 0.48 in our study, and significantly overpredicted *S. griseus* into urban areas where they have not been observed. Therefore, we chose a threshold of 0.6 to more closely reflect the actual distribution of *S. griseus*, as well as to better address our question of

whether particular environmental variables differ in their contribution to the predicted distributions of the three species. We included the 10th percentile training presence logistic threshold for comparison.

We performed linear discriminant function analyses on the output from the model with all variables except land cover (*BioClim No Land cover*) to determine how the environmental variables were associated with the observed western distributions and the predicted western distribution based on models from the native eastern ranges. This model performed as well or nearly as well as the best supported model (see results) and included only continuous variables, which are appropriate for these methods. We used a subset of 10,000 random pixels per species, scaled and normalized the environmental variable data to enable comparison among variables, included 80% in our training sample and 20% as our test sample, and repeated the analyses 1000 times. We ranked the coefficients of linear discriminants by absolute magnitude and compared the rankings among the 1000 replicates to determine the relative variable importance. We then chose one representative replicate (i.e., one that had high model accuracy and followed the trend of variable importance observed across replicates) to visualize and perform a multivariate analysis of variance (MANOVA) to test for significant differences among squirrel taxa.

Results

Species distribution model

Total observations ranged from 2904 to 42,209 per species and geographic region, and 1173 to 16,361 observations after rarefaction (Table 1). In the eastern native extent, the *Full* model was the best supported (highest AUC values) for both *S. carolinensis* and *S. niger* (Table 1). Likewise, for the western extent, we found that the best supported model for *S. griseus* and *S. niger* was the *Full* model (Table 1). However, for *S. carolinensis* in the western extent, the best supported model was *BioClim + Land cover* (Table 1). The *BioClim, No Land cover* model performed equally as well as the *Full* model and nearly as well as the *BioClim + Land cover* model for *S. carolinensis* in the western extent (Table 1). In both western and eastern extents, the model with only BioClim variables was

consistently the lowest supported model for all three species.

For *S. carolinensis* in its native range, BIO12 (mean annual precipitation) had the greatest percent contribution in all models and the highest permutation importance in all but the *BioClim + Land cover* model (Table S1). Probability of occurrence increased with increasing precipitation until approximately 1200 mm, after which there was a slight decline (Figure S1). In the western extent, imperviousness had the highest percent contribution in the *Full* and *BioClim, No Land cover* models; land cover had the highest percent contribution in the *BioClim + Land cover* model; while BIO5 (max temperature of the warmest month) had the greatest permutation importance across all models (Table S1), with greatest probability of occurrence at approximately 25 °C (Figure S1). Probability of occurrence increased with increasing percent developed (imperviousness) until approximately 40%, after which it plateaued (Figure S1). When only BioClim variables were included, BIO6 (minimum temperature of the coldest month) had the greatest percent contribution (Table S1).

For *S. niger* in its native range, imperviousness had the highest percent contribution in the *Full* model and the *No Land cover* model, while mean annual precipitation (BIO12) had the highest permutation importance (Table S1). Probability of occurrence increased with an increase in impervious surfaces and was greatest at an intermediate range of precipitation from 800 to 1000 mm (Figure S2). When imperviousness was excluded (i.e., the *Bioclim + Land cover* model), land cover had the highest percent contribution, with occupancy more associated with lands classified as developed (Table S1). In the western extent, imperviousness had the highest percent contribution for the *Full* and *No Land cover* models while minimum temperature of the coldest month (BIO6) had the highest permutation importance (Table S1). Predicted occupancy increased with increased imperviousness, as well as increased until the minimum temperature of the coldest month reached 9 °C and then plateaued as temperatures increased (Figure S2).

For *S. griseus* in its native range in the western extent, annual precipitation (BIO12) had the highest percent contribution and the highest permutation importance in all four models (Table S1). Probability of occurrence increased with increasing precipitation, reaching an apex at approximately 900 mm, after

which there was a notable decline (Figure S3). The *No Land cover* model performed equally as well as the *Full* model, indicating the NLCD land cover categories did not provide additional information over that contained in the imperviousness and canopy cover variables. Imperviousness had greater percent contribution and permutation importance than canopy cover, with probability of occurrence decreasing at high percent impervious surfaces (Table S1 and Figure S3).

Predicted distribution

The predicted distributions in the western extent for the introduced squirrel species varied considerably depending on whether they were generated from the observations in the eastern versus western extents (Figs. 2 and S5) and there was very little model agreement (Figs. 3 and S6). The novelty score (i.e., how different the combination of variables in the projected extent is from the observed extent) indicated that a high proportion of the western extent was outside the range of observed values in the eastern extent (Figure S4). *Sciurus carolinensis* was predicted to be present in 3.39% of the total area in the western extent based on the eastern model and 0.9% based on the western model (Fig. 2, Table S2). In contrast, *S. niger* was predicted to occur in 0.94% of the total area based on the eastern model and 1.13% based on the western model (Fig. 2, Table S2). For both introduced species, urban areas were largely absent from the predicted distribution generated from the native extent, while they were predicted with high probability from observations in the western extent (Figs. 2 and S6–8). Both models predicted occurrence of *S. carolinensis* in the northern San Francisco Bay Area and Santa Cruz Mountains of California, the Willamette Valley of Oregon, and the Olympia and Tacoma Washington area (Figs. 2 and 3, S6, S8). For *S. niger*, model agreement only occurred around Sacramento and in the Transverse Ranges of Southern California (Figs. 2 and 3, S6, S7), with additional areas of agreement around the San Francisco Bay Area, northern California Central Valley, and Willamette Valley only when considering the lower probability threshold (Figure S6).

There was little predicted geographic overlap among the three squirrel species (Figs. 2 and S5). Based on the model generated from the observations

Fig. 2 Predicted distributions and overlap for all three species in the western United States using a 0.6 probability threshold, based on the best supported models derived from observation in the native extent (**A**) and western extent (**B**)
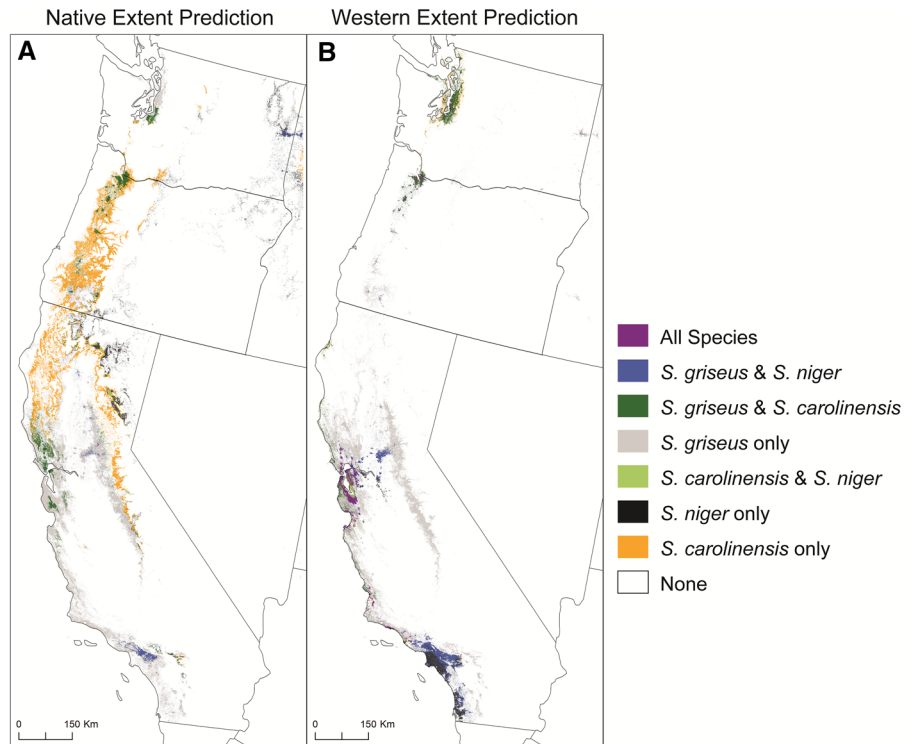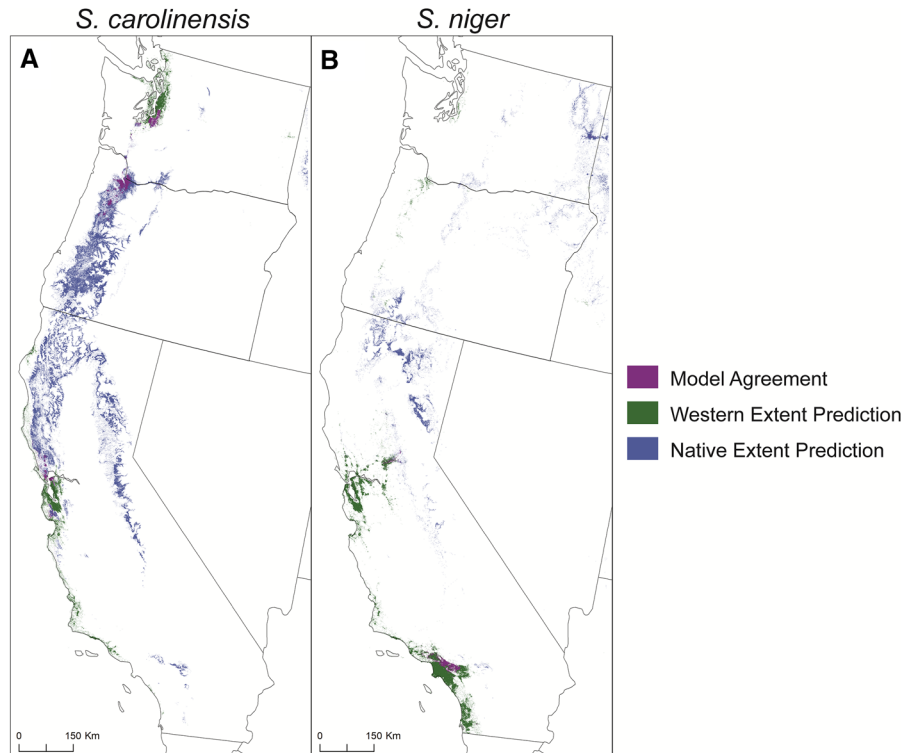

Native Extent Prediction — Western Extent Prediction

Legend:
- All Species
- *S. griseus* & *S. niger*
- *S. griseus* & *S. carolinensis*
- *S. griseus* only
- *S. carolinensis* & *S. niger*
- *S. niger* only
- *S. carolinensis* only
- None

Fig. 3 Map of predicted distributions for *S. carolinensis* (left panels) and *S. niger* (right panels) from the best supported models based on the observations in their native range (blue), their introduced range (green), and where the model predictions agree (purple)

S. carolinensis — *S. niger*

Legend:
- Model Agreement
- Western Extent Prediction
- Native Extent Prediction

in the species' native ranges, all three squirrels are predicted to co-occur in < 1% of the area in the western extent predicted to have at least one squirrel species present, while the western model predicts all species present in 6.7% of areas with at least one squirrel (Fig. 2, Table S2). *Sciurus griseus* was predicted to occur in 3.81% of the total area and 54% of area containing squirrels (Fig. 2, Table S2). Likewise, *S. griseus* was predicted to be the only squirrel present in 77% and 68% of its predicted extent, based on the eastern and western models for the other species, respectively (Table S2). All three species were predicted to overlap considerably within the San Francisco Bay Area based on the western model (Figs. 2 and S5). *Sciurus niger* was the only species predicted to be found in the most densely populated areas of Southern California at the 0.6 threshold (Los Angeles, Orange, and San Diego counties, Fig. 2), while *S.*

griseus and, to a much lesser extent, *S. carolinensis* are predicted to occur there at the lower 10th percentile training presence logistic threshold (Figure S5).

The areas with the highest potential for conflict among the native and introduced species are in the more densely human populated areas. Current conflict with *S. griseus* likely occurs in the western peninsula of the San Francisco Bay Area, Portland, Oregon, and Seattle, Washington areas for *S. carolinensis*, with future conflict predicted primarily adjacent to those areas already occupied (Figs. 4 and 5). Similarly, *S. niger* is likely currently in conflict with *S. griseus* in the San Francisco Bay Area (primarily in the East Bay), Sacramento area, and the Greater Los Angeles Area, and to a lesser extent in Portland Oregon (Figs. 4 and 6). Notably, *S. niger* is predicted to be in conflict with *S. griseus* in northeastern San Diego county, but has not been observed there despite
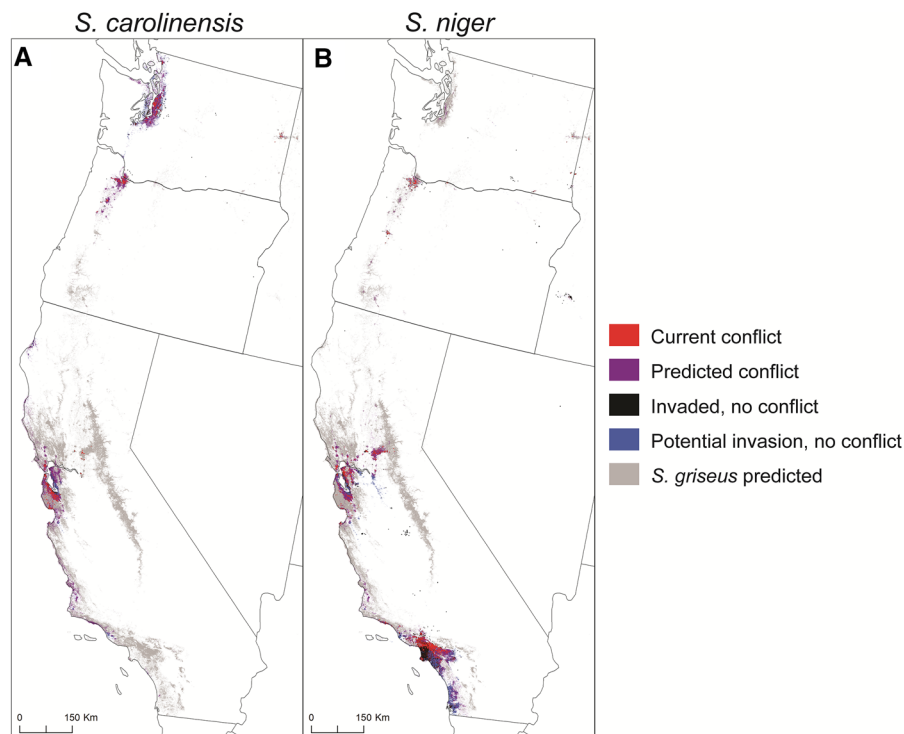

Fig. 4 Current and predicted potential invasion and conflict areas for the native *S. griseus* and introduced *S. carolinensis* (**A**) or *S. niger* (**B**), based on the best supported models in the western extent and iNaturalist observations of introduced squirrels. Current conflict areas are where introduced squirrels have been observed and *S. griseus* is predicted to be (red); predicted conflict areas are where both native and introduced species are predicted to be, but introduced squirrels have not yet been observed (purple); invaded, but no conflict areas are where introduced squirrels have been observed but *S. griseus* is not predicted to be (black); and potential invasion but no conflict are areas where the introduced squirrel is predicted to be but the native squirrel is not (blue)
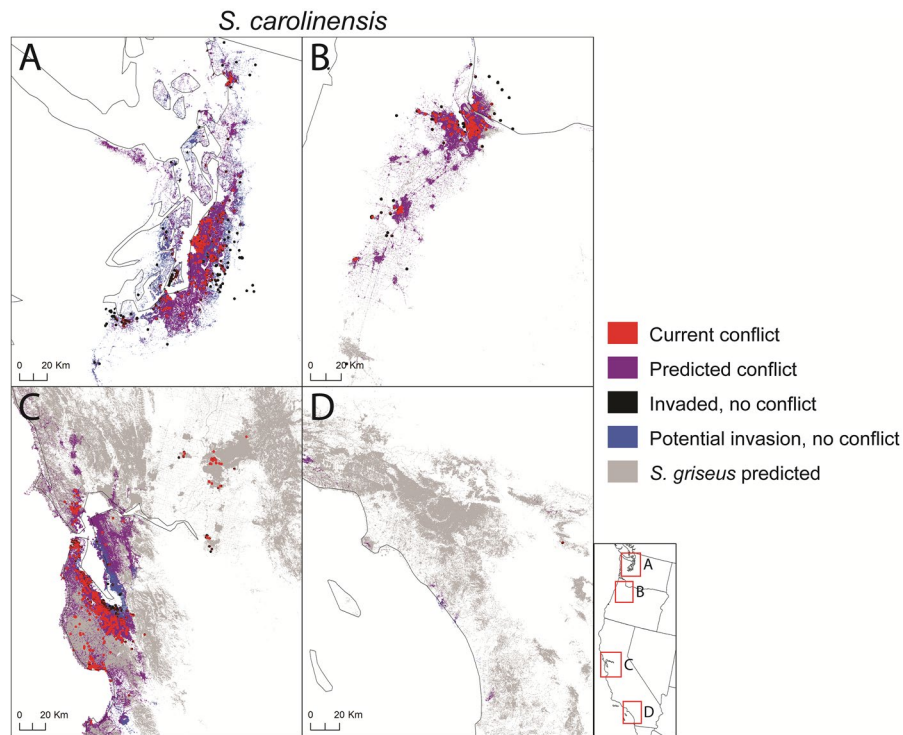
Fig. 5 Current and predicted potential invasion and conflict areas for the native *S. griseus* and introduced *S. carolinensis* in the Seattle (**A**), Portland (**B**), San Francisco Bay Area (**C**), and Los Angeles Metropolitan area south to San Diego county (**D**), based on the best supported models in the western extent and iNaturalist observations of *S. carolinensis*. Current conflict areas are where introduced squirrels have been observed and *S. griseus* is predicted to be (red); predicted conflict areas are where both native and introduced species are predicted to be, but introduced squirrels have not yet been observed (purple); invaded, but no conflict areas are where introduced squirrels have been observed but *S. griseus* is not predicted to be (black); and potential invasion but no conflict are areas where the introduced squirrel is predicted to be but the native squirrel is not (blue)

observations in southwestern San Diego (Fig. 6). Likewise, both introduced species are predicted to occur along the central California coast in San Luis Obispo County, but neither has been observed there (Fig. 4). Predicted conflict areas increase to include much of the California coastal areas for *S. carolinensis* and the Southern California counties of Santa Barbara, Los Angeles, Orange, and San Diego for *S. niger* when considering the lower presence threshold (Figures S9-S11).

Discriminant function analysis

The accuracy in assigning test data to the correct species ranged from 0.72 to 0.76 (mean 0.74) for the species distribution models derived from the native ranges projected to the western extent, and 0.70 to 0.74 (mean 0.72) for the models derived

from the western observations only. There was a statistically significant difference among species on the combined dependent variables for the predicted distributions based on the eastern range (F(12, 47984)=3018.8, p<0.0001) and western range (F(12, 47984)=2431.3, p<0.0001). The variables with the greatest magnitude effect in discriminating among the three taxa were annual precipitation (BIO12), minimum temperature of the coldest month (BIO6), and annual mean temperature (BIO1) for the models projected from the native ranges. For the models derived from the western extent, minimum temperature of the coldest month, max temperature of warmest month (BIO5), and imperviousness had the greatest effect (Fig. 7, Tables S3 and S4). Results were similar with the lower 10th percentile training presence logistic threshold, except minimum temperature was less discriminating in the native range
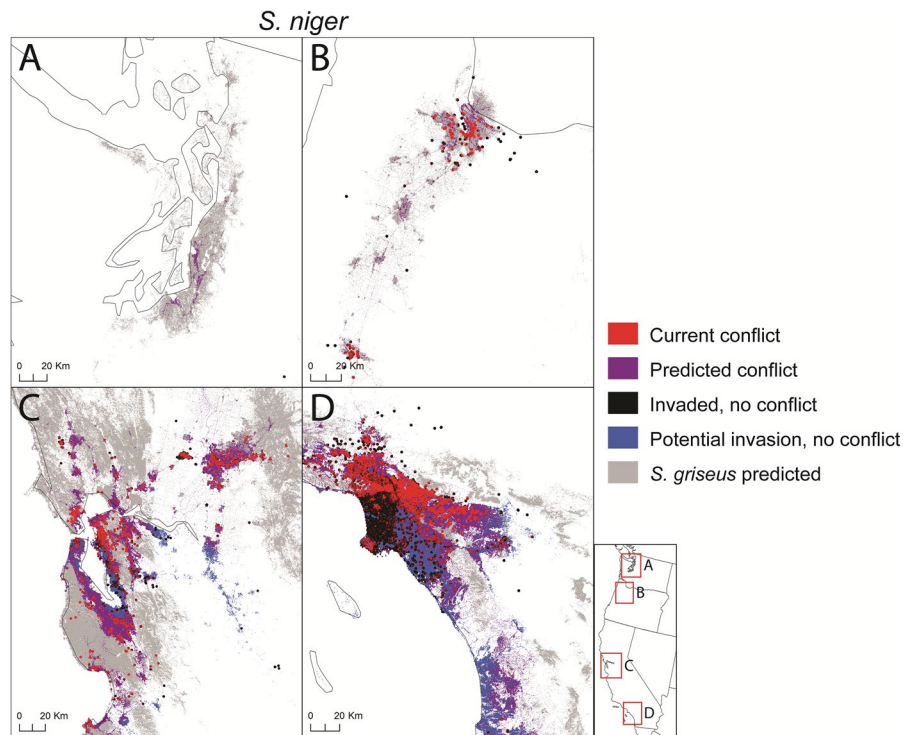
Fig. 6 Current and predicted potential invasion and conflict areas for the native *S. griseus* and introduced *S. niger* in the Seattle (**A**), Portland (**B**), San Francisco Bay Area (**C**), and Los Angeles Metropolitan area south to San Diego county (**D**), based on the best supported models in the western extent and iNaturalist observations of introduced squirrels. Current conflict areas are where introduced squirrels have been observed and *S. griseus* is predicted to be (red); predicted conflict areas

are where both native and introduced species are predicted to be, but introduced squirrels have not yet been observed (purple); invaded, but no conflict areas are where introduced squirrels have been observed but *S. griseus* is not predicted to be (black); and potential invasion but no conflict are areas where the introduced squirrel is predicted to be but the native squirrel is not (blue)

Fig. 7 Scatterplot of the discriminant function analysis of the environmental variables obtained from the three predicted species distributions for the models derived from the native ranges (left) and western extent (right)
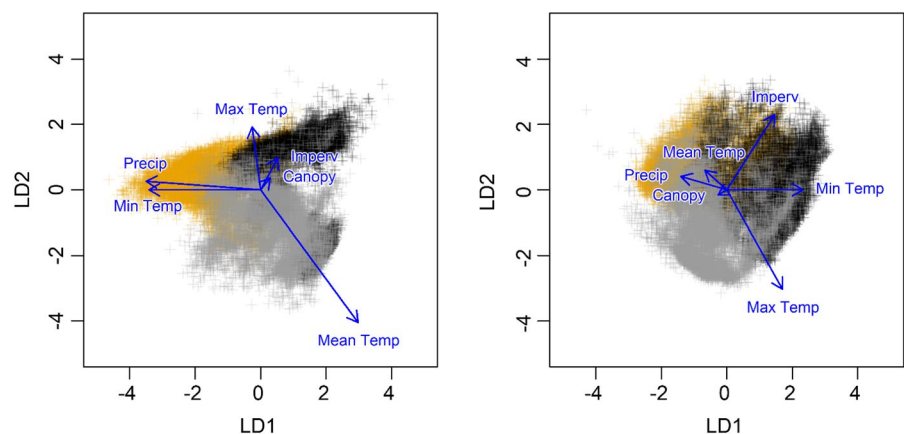
predictions, and imperviousness less discriminating in the western extent predictions (Table S4). There was greater overlap among the three species in the

model derived from the western extent. *Sciurus carolinensis* was predicted to occupy wetter areas and *S. griseus* to occupy areas with warmer mean annual
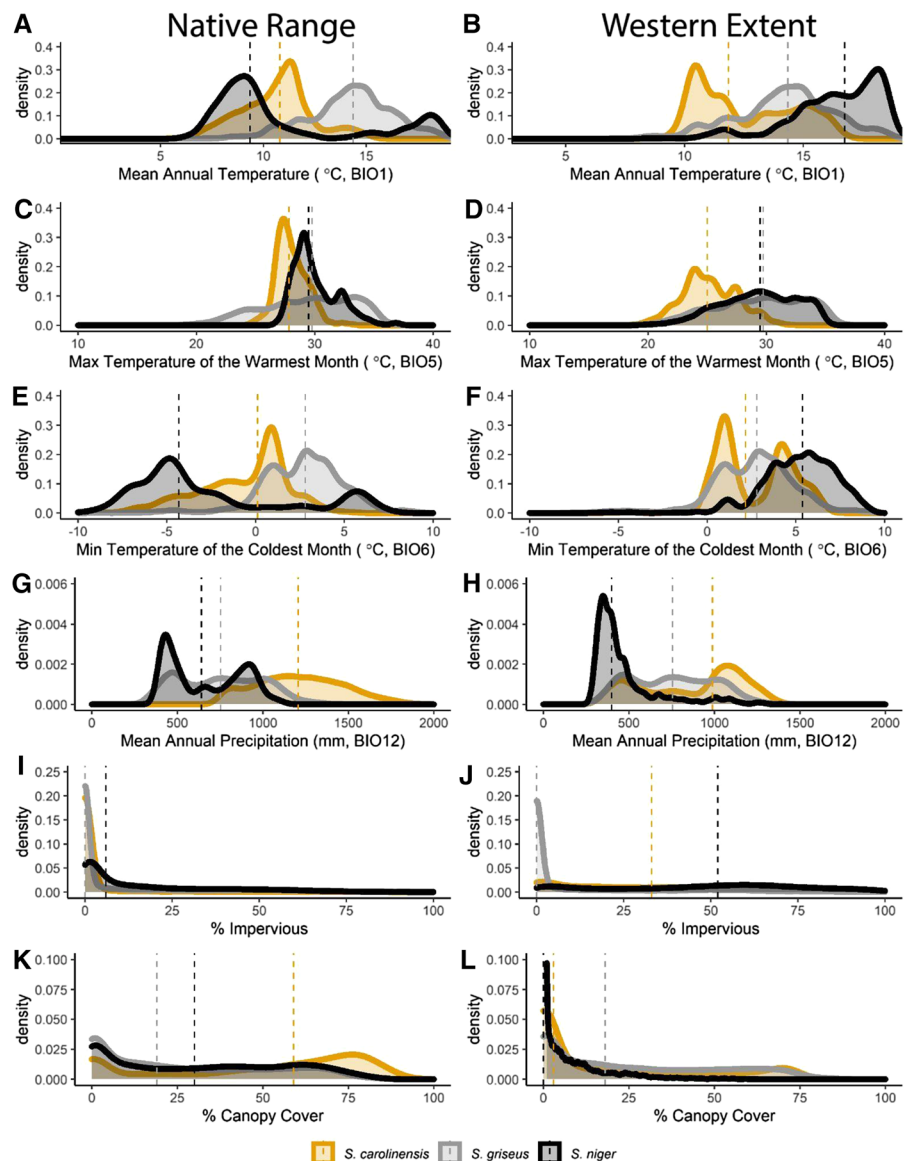
temperatures in the models derived from the native extents (Figs. 7 and 8, S12). The models derived from the western extent predicted *S. carolinensis* to occupy colder areas than the other two species and *S. griseus* was predicted to occupy areas with less imperviousness (Figs. 7 and 8, S12). Percent canopy cover was higher for the introduced squirrels in their native range than *S. griseus*, despite *S. griseus* being restricted to forested areas, and higher than predicted in the western extent (Fig. 8). This is most likely because percent canopy cover values were generally higher in the eastern extent than western extent and

highest in the deciduous and mixed forests of the eastern United States.

Discussion

Abundant observation data collected by members of the general public on the iNaturalist platform enabled a thorough analysis of the current invasion status of two eastern squirrel species and their potential impacts on the native western grey squirrel. Specifically, we found that expansion of the non-native

Fig. 8 Density plots of environmental variables for each of the three predicted squirrel species distributions based on the model derived from observations in their native ranges (left column) and the western range (right column). Vertical lines are median values

squirrels is predicted by both the models derived from the native and western extent observations, but with different predicted distributions. Potential conflict with the native western grey squirrel is primarily predicted to occur in and around areas of human development. Imperviousness was both an important predictor of the introduced squirrels and a discriminating factor between the introduced and native species, such that *S. griseus* is predicted to occupy areas less impacted by humans.

The predicted distributions of non-native squirrels in the western United States differed considerably depending on whether the model was built using observations in the native or introduced range, however, both sets of models predicted a potential expansion of the introduced species in the west coast states. For *S. carolinensis*, there is potential for expansion into the mountainous areas of Oregon and northern and central California, based on the model derived from the native extent. Expansion into southern California is limited to coastal regions and primarily predicted by the western extent observations. Particularly noteworthy is the potential for expansion around the densely populated areas of the San Francisco Bay Area, Portland, and Seattle, which were strongly predicted by both sets of models. Interestingly, despite there being many more observations of *S. niger* over a broader geographic area, the predicted potential for expansion is less than that of *S. carolinensis* and is primarily found in and around urban areas.

The areas where the introduced squirrels are predicted to overlap most with *S. griseus* are those areas adjacent to areas with high human density, such as the foothills, mountains and forested areas surrounding Los Angeles and the San Francisco Bay Area. Therefore, as human encroachment into wildlands continues, the modeling presented here projects increased contact, and in turn conflict, among these three species. There is some evidence that *S. carolinensis* is displacing the native squirrel in the Santa Cruz Mountains (Jessen et al. 2018). Jessen et al. (2018) note that *S. carolinensis* can competitively exclude *S. griseus* in developed areas and fragmented forests. Likewise, *S. niger* has been documented replacing *S. griseus* in Southern California (Muchlinski et al. 2009) and is predicted to continue displacing the native tree squirrel (Garcia and Muchlinski 2017). Additionally, *S. niger* and *S. carolinensis* co-occur throughout much of their native range, potentially providing a competitive advantage over *S. griseus* which has not historically had to compete with other large arboreal squirrel species.

Sciurus griseus is already negatively impacted by urbanization and fragmentation (Jessen et al. 2018; DeMarco et al. 2021) and these effects may be amplified by the presence of introduced tree squirrels. Importantly, if the non-native squirrels are limited by dispersal, continuing development is likely to facilitate their spread into previously inaccessible forested areas, increasing potential displacement of the native squirrel. Our results suggest that there is high potential for *S. carolinensis* to expand its range in the coastal states into much of the forested areas as predicted by the eastern extent model, while *S. niger* is likely restricted to urban areas, suburban areas, and areas immediately surrounding them.

Of our four sets of models, the *BioClim Only* model performed the worst in every species and extent, indicating that vegetation and especially human impact variables, are critical for understanding the invasion trajectories for these three squirrels. The *BioClim, No Land cover* model performed similarly or better than the *BioClim + Land cover* model, while imperviousness had greater percent contribution and permutation importance than land cover and canopy cover, suggesting that degree of urbanization is a better predictor of squirrel distributions than categorical habitat type in both extents. For the two introduced species, human impact had a greater contribution to the species distribution models in the western extent than the eastern extent, while different bioclimatic variables were more important in the different extents. This latter observation is not surprising given the different extents have very different environmental conditions, as evidenced by the low novelty scores produced in Maxent (i.e., the combination of variables contributing to the distribution in the native extent differed from those found in much of the western extent; Figure S4). *Sciurus niger* is predicted to occur in much hotter areas in the western extent than the native extent, while both introduced species are predicted to occur in drier areas and areas with more impervious surfaces. Likewise, *S. griseus* is predominantly predicted to be in areas with < 5% impervious surfaces, while *S. carolinensis* and *S. niger* are predicted to occur over a wide range of imperviousness, indicating a greater ability to persist in anthropogenic landscapes than the native squirrel.

The current distribution of introduced squirrels in the western extent strongly reflects the introduction history of the two species. For example, both species were introduced into the San Francisco Bay Area, but only *S. niger* has been successfully introduced into Southern California (Byrne 1979). Currently, *S. carolinensis* is the most abundant species in the San Francisco peninsula (Mountain View to South San Francisco), *S. niger* is the most abundant in the East Bay, and *S. griseus* is the most abundant in the Santa Cruz Mountains, while both introduced species co-occur in San Jose and San Francisco, and all three species occur in Marin County. The predicted distribution from the western extent in turn is a much better fit to the observed distribution, thus the disparity between the predictions from the eastern versus western extent is also likely explained, at least in part, by the introduction history of the squirrels. However, the mechanisms underlying the incongruence remain unknown. One possible explanation is a founder effect and subsequent selection favoring individuals that are more tolerant of humans. Likewise, if their distributions were limited primarily by dispersal, we would expect areas with high modeled suitability but no known occurrences separated from areas with known occurrences by regions with low suitability. We see examples of this in both species in the central California coast.

Our results are similar to previous studies that have looked at the two introduced species individually, but with some notable differences. Garcia and Muchlinski (2017) modeled the distribution of *S. niger* in the Greater Los Angeles Metropolitan Area. Creley et al. (2019) modeled the distribution of *S. carolinensis* throughout California. Both studies used a similar approach as the present study, but included 19 Bio-Clim variables as well as elevation, which was not included in this study. The authors also combined observations from the native ranges with observations in the Greater Los Angeles Metropolitan Area (*S. niger*) and throughout California (*S. carolinensis*) to generate their species distribution models. Despite these differences, the predicted distributions of *S. niger* in Southern California are very similar. Our predictions derived from the western observations are nearly identical, with slightly more suitable habitat predicted in the foothills of the San Gabriel Mountains. This may reflect movement of *S. niger* into these areas between their data collection in 2015

and ours in 2020. Approximately 20 observations of *S. niger* have been made in the San Gabriel Mountains since 2017, with very few prior to then. Muchlinski et al. (2012) suggested that ground cover height and percent canopy cover, particularly oak tree cover, was an important predictor of coexistence, exclusion of *S. niger* by *S. griseus* (highest percent oak trees), or replacement of *S. griseus* by *S. niger* (lowest percent oak trees) in Southern California. Based on these results, they predicted that *S. niger* might soon move into the San Gabriel, San Bernardino, and San Jacinto Mountains, which is unfortunately supported by our study. Our model derived from the eastern extent observations predicts more suitable habitat in the San Gabriel and San Bernardino Mountains than the previous study (Garcia and Muchlinski 2017 Fig. 3), however this is only true at the lower threshold (Figure S7).

The previous study of *S. carolinensis* predicted significantly more suitable habitat in Southern California, but less area overall when comparing model outputs at the same 10th percentile training presence logistic threshold, particularly for the model derived from the native range (Figure S6 and Creley et al. 2019 Fig. 1). This discrepancy in Southern California is likely due to the inclusion of observations from Oregon and Washington in the present study that were not included in the previous study, which could shift the predicted distribution for *S. carolinensis* toward cooler and wetter areas and away from warmer and drier Southern California. Our *S. niger* observations ranged from extreme southern California to Portland, Oregon with a few observations throughout Washington, while our *S. carolinensis* observations are abundant in the Pacific Northwest and do not extend further south than Monterey County, California. Therefore, it is not surprising that our predicted distribution for *S. niger* more closely matches those of the previous study than does *S. carolinensis*. This underscores the impact that sampling extent has on model output and should be considered when designing studies.

Our analyses are based entirely on observations made by community scientists. While this provides us with an enormous amount of valuable data, there are limitations to working with these types of records. We restricted all observations to "research grade" which have been verified by multiple observers, but errors are still possible, especially given the morphological

variation within both *S. niger* and *S. carolinensis* that make accurate identification challenging for even expert observers. While it is not possible to manually verify the nearly 75,500 observations, the number of incorrectly identified squirrels is likely very low compared to the total number of observations. Additionally, observations are biased towards urban, suburban, and otherwise easily accessible areas where people are more likely to go. While this is certainly a limitation of the data, the distribution of observations throughout the study area, in particular the *S. griseus* observations in the mountainous and rural areas of California, suggests that the data are not biased towards one species and they are sufficiently well distributed to accurately reflect the actual distribution of the three species. Maxent is particularly well-suited to handle these types of imperfect data. Finally, we chose to use the most recently published BioClim data because it is at an appropriate spatial extent and scale for our analyses, as well as to be consistent with previous studies. Future studies could use more recent data, and or average over multiple time periods.

This is the first study to simultaneously model species distributions for these three squirrel species in their native and introduced ranges. Our results indicate high potential for the non-native species to expand their ranges within the coastal states as a result of increased human footprint and potentially climate change as conditions become less favorable for *S. griseus*. The most conservative estimates, derived from observations in the western extent and at the higher 0.6 threshold, suggest limited expansion beyond human-dominated landscapes, while the predictions from both models at the lower threshold suggest the two introduced species could have significant negative impacts on *S. griseus*. Citizen science databases, such as iNaturalist, are valuable sources of data that will enable us to continue to track the spread of introduced species with much greater temporal and spatial resolution, which can provide us with a better understanding of the mechanisms enabling invasions. Future research should monitor the continued spread of these squirrels and determine whether the introduced squirrels are able to move into the areas predicted to be suitable based on their native range but not their current distribution, or if they are restricted to areas of high anthropogenic presence.

References

Bertolino S (2009) Animal trade and non-indigenous species introduction: the world-wide spread of squirrels. Divers Distrib 15:701–708

Bertolino S, Genovesi P (2005) The application of the European strategy on invasive alien species: an example with introduced squirrels Hystrix. Ital J Mammal. https://doi.org/10.4404/hystrix-16.1-4343

Bertolino S, Currado I, Mazzoglio P, Amori G (2000) Native and alien squirrels in Italy Hystrix. Ital J Mammal 11:65–74

Boria RA, Olson LE, Goodman SM, Anderson RP (2014) Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. Ecol Model 275:73–77

Brown JL (2014) SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. Methods Ecol Evol 5:694–700

Byrne S (1979) The distribution and ecology of the non-native tree squirrels Sciurus carolinensis and Sciurus niger in northern California. University of California, Berkeley

Carraway LN, Verts BJ (1994) Sciurus griseus. Mamm Species. https://doi.org/10.2307/3504097

Claytor SC, Muchlinski AE, Torres E (2015) Multiple introductions of the eastern fox squirrel (*Sciurus niger*) in California. Mitochondrial DNA 26:583–592

Cooper DS, Muchlinski AE (2015) Recent decline of lowland populations of the western gray squirrel in the Los Angeles area of Southern California. Bull South California Acad Sci 114(42–53):12

Creley CM, Shilling FM, Muchlinski AE (2019) An Ecological niche model to predict range expansion of the Eastern gray squirrel in California. Bull South California Acad Sci 118(58–70):1

DeMarco C, Cooper DS, Torres E, Muchlinski A, Aguilar A (2021) Effects of urbanization on population genetic structure of western gray squirrels. Conserv Genet 22:67–81

Dewitz J (2019) National land cover database (NLCD) 2016 products: U.S. Geological Survey data release. https://doi.org/10.5066/P96HHBIE.

Di Febbraro M, Martinoli A, Russo D, Preatoni D, Bertolino S (2016) Modelling the effects of climate change on the risk of invasion by alien squirrels Hystrix. Ital J Mammal 27:1–8

Garcia RB, Muchlinski AE (2017) Range expansion of the Eastern fox squirrel within the greater Los Angeles metropolitan area (2005–2014) and projections for continued range expansion. Bull South California Acad Sci 116(33–45):13

Hellmann JJ, Byers JE, Bierwagen BG, Dukes JS (2008) Five potential consequences of climate change for invasive species. Conserv Biol 22:534–543

Jessen T, Wang Y, Wilmers CC (2018) Habitat fragmentation provides a competitive advantage to an invasive tree squirrel, Sciurus carolinensis. Biol Invasions 20:607–618

Koprowski JL (1994) Sciurus carolinensis. Mamm Species. https://doi.org/10.2307/3504224

Koprowski JL (1994) Sciurus niger. Mammal Species. https://doi.org/10.2307/3504263

Martinoli A, Bertolino S, Preatoni DG, Balduzzi A, Marsan A, Genovesi P, Tosi G, Wauters LA (2010) Headcount 2010: the multiplication of the grey squirrel populations introduced to Italy Hystrix. Ital J Mammal 21:127–136

Mccleery RA, Lopez RR, Silvy NJ, Gallant DL (2008) Fox squirrel survival in urban and rural environments. J Wildl Manag 72(133–137):135

Merow C, Smith MJ, Silander JAJ (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography 36:1058–1069

Muchlinski AE, Stewart GR, King JL, Lewis SA (2009) Documentation of replacement of native Western gray squirrels by introduced Eastern fox squirrels. Bull South California Acad Sci 108(160–162):163

Muchlinski A, Gatza B, Lewis S, Erkebaeva K (2012) Eastern fox squirrels and western gray squirrels in southern California. Proc Vertebr Pest Conf. https://doi.org/10.5070/V425110358

Ortiz JL (2022) Temporal and spatial overlap in the behaviors of a native and invasive tree squirrel in southern California. Ethol Ecol Evol 34(148–164):2

Ortiz JL, Muchlinski AE (2015) Food selection of coexisting Western gray squirrels and Eastern fox squirrels in a native California botanic garden in Claremont, California. Bull South California Acad Sci 114(98–103):106

Parker TS, Nilon CH (2008) Gray squirrel density, habitat suitability, and behavior in urban parks. Urban Ecosyst 11:243–255

Pearson RG (2010) Species' distribution modeling for conservation educators and practitioners. Lessons Conserv 3:54–89

Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31:161–175

Phillips SJ, Anderson RP, Dudík M, Schapire RE, Blair ME (2017) Opening the black box: an open-source release of Maxent. Ecography 40:887–893

Phillips SJ, Dudík M, Schapire RE Maxent software for modeling species niches and distributions (Version 3.4.1). Available from url: http://biodiversityinformatics.amnh.org/open_source/maxent/

Rödder D, Engler JO (2011) Quantitative metrics of overlaps in Grinnellian niches: advances and possible drawbacks: quantitative metrics of niche overlap. Global Ecol Biogeogr 20(6):915–927. https://doi.org/10.1111/j.1466-8238.2011.00659.x

Sala OE, Stuart Chapin FI, Armesto JJ, Berlow E, Bloomfield J, Dirzo R, Huber-Sanwald E, Huenneke LF, Jackson RB, Kinzig A, Leemans R (2000) Global biodiversity scenarios for the year 2100. Science 287:1770–1774

Team R C (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Twining JP, Montgomery WI, Tosh DG (2020) The dynamics of pine marten predation on red and grey squirrels. Mamm Biol 100:285–293

Venables WN, Ripley BD (2002) Modern applied statistics with S, 4th edn. Springer, New York

Wauters LA, Tosi G, Gurnell J (2002) Interspecific competition in tree squirrels: do introduced grey squirrels (Sciurus carolinensis) deplete tree seeds hoarded by red squirrels (S. vulgaris)? Behav Ecol Sociobiol 51:360–367

Williamson RD (1983) Identification of urban habitat components which affect eastern gray squirrel abundance. Urban Ecol 7:345–356