

# Substitutable habitats? The biophysical and anthropogenic drivers of an exotic bird's distribution

Amélie Y. Davis · Nur Malas · Emily S. Minor

Received: 11 January 2013 / Accepted: 25 July 2013 / Published online: 4 August 2013  
© Springer Science+Business Media Dordrecht 2013

**Abstract** The spread and distribution of exotic species depends on a number of factors, both anthropogenic and biophysical. The importance of each factor may vary geographically, making it difficult to predict where a species will spread. In this paper, we examine the factors that influence the distribution of monk parakeets (*Myiopsitta monachus*), a parrot native to South America that has become established in the United States. We use monk parakeet observations gathered from citizen-science datasets to inform a series of random forest models that examine the relative importance of biophysical and anthropogenic variables in different regions of the United States. We find that while the distribution of monk parakeets in the southern US is best explained by biophysical variables such as January dew point temperature and forest cover, the distribution of monk parakeets in the northern US appears to be limited to urban environments. Our results suggest that monk parakeets are unlikely to spread outside of urban environments in the northern United States, as they are not adapted to the climatic conditions in that region. We extend the

notion of “substitutable habitats,” previously applied to different habitats in the same landscape, to exotic species in novel landscapes (e.g., cities). These novel landscapes provide resources and environmental conditions that, although very different from the species' native habitat, still enable them to become established. Our results highlight the importance of understanding the regionally-specific factors that allow an exotic species to become established, which is key to predicting their expansion beyond areas of introduction.

**Keywords** Human–environment interaction · Variable importance · Species establishment · Habitat suitability modeling · Introduced species · Bird feeders

## Introduction

The distribution of an exotic species is the outcome of complex interactions between the species' biology, the environment into which the species is introduced, and anthropogenic disturbances that create novel ecological opportunities (Thuiller et al. 2006). Because urban landscapes provide such ecological opportunities and are often the point of introduction for new species, many exotic species thrive in cities (McKinney 2006). Predicting which species will go beyond urban areas and where they will spread is important for managing exotic species and preserving biodiversity.

---

A. Y. Davis · E. S. Minor  
Institute for Environmental Science and Policy,  
University of Illinois at Chicago, 2121 West Taylor Street  
(MC 673), Chicago, IL 60612, USA

N. Malas · E. S. Minor (✉)  
Biological Sciences, University of Illinois at Chicago, 845  
W. Taylor Street (MC 066), Chicago, IL 60607, USA  
e-mail: eminor@uic.edu

One factor that is frequently used to explain patterns of invasion is the suitability of climate or habitat in the new environment (Blackburn and Duncan 2001; Moyle and Marchetti 2006; Hulme 2009; Murray et al. 2012). As Sexton et al. (2009) state, “species range limits are essentially the expression of a species’ ecological niche in space”. If environmental conditions in an area are unsuitable for a species, the probability of its establishment at that location is low. However, even in suitable environmental conditions, a species introduction may not lead to a viable population due to stochastic die-off events (Simberloff 2009). Other factors that affect establishment include natural enemies, competitors, and resource availability (Shea and Chesson 2002; Tilman 2004; MacLeod et al. 2009). The temporal and spatial patterns of propagule pressure are also thought to play an important role in establishment success (Catford et al. 2009; Chiron et al. 2009; Simberloff 2009). Propagule size, or the number of individuals in a release event, can act to reduce the effect of demographic stochasticity, while propagule number, or the number of distinct release events, can reduce the impact of environmental stochasticity (Simberloff 2009). Greater propagule pressure also increases the chance of a successful invasion by increasing genetic diversity and continually supplementing the population (Colautti et al. 2006).

Because habitat suitability is thought to be so important in the establishment and subsequent spread of an invasive species, modeling habitat suitability for exotic species is an active and important area of research (Thuiller et al. 2005). Habitat suitability is often measured in terms of biophysical variables such as temperature, precipitation, or vegetation type, in an attempt to identify areas with climates similar to the native range (termed “climate matching” (Peterson 2003; Bomford et al. 2009)). But defining habitat suitability may go beyond climate matching, and biophysical variables alone might not always be adequate for explaining the spread and distribution of exotic species. Recent studies have highlighted the importance of human-related factors as well as biophysical factors in shaping the distribution of exotic species (Roura-Pascual et al. 2011; Pyšek et al. 2010). In fact, Roura-Pascual et al. (2011) showed that, by creating favorable microclimates, human modification of environments may allow exotic species such as the Argentine ant to establish in places where climatic suitability is low.

Urban areas in particular may present unique habitat types for exotic species. They contain anthropogenic food sources, which are often abundant and consistently available (Rebele 1994; Lepczyk et al. 2004). This is compounded by the fact that exotic species are often adept at utilizing resources which native species might not (Sol et al. 2012) and are likely to have high rates of feeding innovation (Møller 2009). Furthermore, native species tend to be scarce in urban areas (vanHeezik et al. 2008), reducing biotic resistance (Levine et al. 2004) and creating empty ecological niches that can potentially be filled by opportunistic exotic species (Sol et al. 2012). Urban areas usually differ climatically and physically from the surrounding land (Arnfield 2003), creating microclimates or habitats that may be more suitable for some species than less-disturbed outlying areas (Sukopp 2004; Song et al. 2012). Finally, propagule pressure may be more intense in urban areas (Chytrý et al. 2008). Both the altered environment and increased propagule pressure may allow exotic species to persist in a region that might otherwise be unsuitable.

If persistence of an exotic species in a particular region is dependent on conditions created by humans, human-related factors (e.g., housing density) will be important determinants for predicting presence of that species. However, those same factors may have minimal importance in regions that are already highly suitable for the same species. If human-related factors are important in some environments and not in others, this could introduce error into habitat suitability models or make their output difficult to interpret. These problems could be magnified if analyses are conducted over large spatial scales where human and environmental factors vary greatly. By carving out different subgroups of a species based on geographic location and developing habitat suitability models for each, we can begin to understand what factors are important under different conditions. This may be key to predicting and ultimately preventing or limiting the spread of exotic species.

In this paper, we disentangle the factors that affect the distribution of an exotic bird, the monk parakeet (*Myiopsitta monachus*). Specifically, we examine the relative importance of biophysical variables and anthropogenic factors on monk parakeet distribution across the contiguous United States. We conducted our analysis for different geographic regions, with the

expectation that the importance of each factor would depend on the scale and region of analysis. Specifically, we predicted that human-related factors would be more important in northern regions, where monk parakeets seem to rely on bird feeders in the winter months (Hyman and Pruett-Jones 1995). Our results highlight the importance of running species distribution models at different spatial extents and geographic locations.

## Methods

### Focal species

Monk parakeets are medium-sized parrots native to South America (Spreyer and Bucher 1998). In their native range, they are considered an agricultural pest (Lever 2006). Like other members of the parrot family, they are highly gregarious and often nest communally. Unlike other members of the parrot family, they build stick nests and use them year-round (Spreyer and Bucher 1998). They are common birds in the pet trade and, as a result of escapes or releases, have established non-native populations throughout the United States and the world (Strubbe and Matthysen 2009; Russello et al. 2008; Van Bael and Pruett-Jones 1996). Their spread has elicited concern about their potential impact on agricultural crops in their introduced range (Davis 1974). While they are found in some climates that are colder than their native range, it has been suggested that monk parakeets are reliant on bird feeders in these locations (Hyman and Pruett-Jones 1995). In this paper, we examine the factors that influence their distribution in the contiguous United States.

### Bird observations

We used monk parakeet observations gathered from citizen-science datasets to inform our model and contrasted these observations with a set of pseudo-absence locations. To decrease the probability that monk parakeet absence points were simply locations with few observers (i.e., false absences), and to reduce the effect of sample bias on our results, we used reported observations of a common and widely-distributed native species, the American robin (*Turdus migratorius*), as pseudo-absence or background

points. By using a common species, we created a widely-distributed set of background points in locations where volunteers go birding, excluding from the model any locations that are too remote or inaccessible for us to have any knowledge about the birds that are found there. This set of background points therefore represents the inherent spatial bias in the monk parakeet observations reported through the citizen science datasets. Using this approach ensures that the environmental conditions associated with the background points contain the same spatial bias as the environmental conditions associated with the presence points (Phillips et al. 2009; Mateo et al. 2010).

Monk parakeet and American robin observations were obtained from four citizen science programs: eBird, Project Feederwatch (PWF), Great Backyard Bird Count (GBBC) and the Christmas Bird Count (CBC). Each program has a different approach to collecting bird observations. The National Audubon Society runs the CBC (<http://birds.audubon.org/christmas-bird-count>) from mid-December to early January each year. Volunteers count every bird seen or heard in a single day along specific routes within a “count circle” with a 15 mile (24 km) diameter. The Cornell Laboratory of Ornithology runs PWF and GBBC. For PWF (<http://www.birds.cornell.edu/pfw/>), volunteers pick an area where they will count birds, such as their backyard; most observation areas are about the size of two tennis courts. Volunteers count each bird they see for two consecutive days, counting at any time of day and for as little time as they want (effort is recorded). This observation period can occur at any time during the 21-week period starting on the second Saturday in November. The GBBC is a four-day event in February, before spring migration (<http://www.birdsource.org/gbbc>). Volunteers count every bird they observe during 15 min (minimum) at any time (even multiple times) throughout the 4-day period. The eBird program is a little different, as it has fewer restrictions than the other three programs (<http://eBird.org>). eBird is a collaboration between the Cornell Laboratory of Ornithology and the National Audubon Society. Citizen-scientists are free to report any birds they see at any time and location. Some volunteers do traveling counts and report the distance they hiked, others enter data for stationary counts. Other options include casual sightings and exhaustive area counts.

All these citizen datasets are biased in terms of spatial and temporal coverage (Dickinson et al. 2010;

Yaukey 2010) and they oversample in places where more people live (Sullivan et al. 2009). Some programs have more standardized protocols than others (Yaukey 2010) and all of them are subject to errors in identification and detectability (Sullivan et al. 2009). However, based on expert lists of birds in an area potentially faulty sightings are flagged, and regional experts determine whether they should be removed from the database (Sullivan et al. 2009). In general, the information in these datasets tends to be accurate and reliable (Munson et al. 2010; Bonter and Cooper 2012).

We used observations from ebird, PWF, GBBC, and CBC reported between January 2007 and December 2011, and downloaded from <http://www.avianknowledge.net>. The observations were georeferenced in ArcGIS10.0 (ESRI 2011). The GIS analyses were conducted in ArcGIS 10.0 and 9.3. We divided the contiguous United States into  $10 \times 10$  km grid cells (Albers Equal Area Conic Projection) and retained all grid cells with either monk parakeet or American robin observations for analysis. The grid was created in Hawth's Analysis Tools v3.27 (Beyer 2004). Any grid cell containing a monk parakeet observation, whether or not it also contained an American robin observation, was considered a monk parakeet "presence" point. Grid cells with only American robin observations were considered monk parakeet "absence" points, i.e., pseudo-absences.

### Predictor variables

Our choice of predictor variables was guided by a literature search and knowledge of monk parakeet physiology, diet, and native range bioclimatic characteristics. In their native range, monk parakeets are known to inhabit open woodlands and scrublands and are also crop pests (Lever 2006). Distance to cities of 100,000 and 500,000 inhabitants, distance to highways, human population density, slope of terrain, number of frost days, temperature, precipitation, and mean normalized difference vegetation index (NDVI), have all been found to be important to monk parakeet presence in Europe (Muñoz and Real 2006; Strubbe and Matthysen 2009). In Barcelona, Spain, monk parakeets tend to avoid areas that are heavily forested (Sol et al. 1997) and in Chicago, IL (USA), distance to railways was an important predictor of monk parakeet nest locations (Minor et al. 2012). Finally, as monk

parakeets are known to visit bird feeders, and household income has been related to the provision of bird feeders (Fuller et al. 2008), this might be a relevant variable explaining their distribution. Therefore, we measured the following variables for all grid cells with American robin or monk parakeet presences: number of frost days; January dew point temperature; precipitation; NDVI; slope; percent of land covered by forest, scrub, and crops; housing density; household income; and distance to the nearest railway, interstate highway, town with 100,000 inhabitants, and town with 500,000 inhabitants (Table 1).

The environmental variables were derived from a number of different sources (Table 1). We used zonal statistics in ArcGIS 9.3 to determine the mean value of each variable for each  $10 \times 10$  km grid cell used in the model. For linear features (i.e., roads and railways) and points (i.e., towns), we first used the Euclidean distance tool in ArcGIS to calculate the distance from  $1,000 \times 1,000$  m grid cells to the nearest feature of interest. We then used zonal statistics to average these distances for each  $10 \times 10$  km grid cell. Housing density and median household income were derived from polygons (US Census block groups and census tracts, respectively) and then summarized at the level of each  $10 \times 10$  km grid cell using the "polygon in polygon analysis" in Hawth's Analysis Tools (Beyer 2004).

Dormann et al. (2012) suggest that strongly correlated variables ( $\rho \geq 0.7$ ) be removed from subsequent modeling, so we excluded winter NDVI and mean January temperature from our analysis. All remaining variables had correlation coefficients less than 0.7 and were subsequently used as predictor variables in our species distribution models.

### Species distribution modeling

We modeled species distribution of monk parakeets using random forests (Breiman 2001). Random forests are an ensemble learning method that aggregates multiple classification and regression trees (CART models) into a single, more accurate, predictive model. Trees are created from bootstrap samples of the data, and a random subset of the predictor variables are used to split each node of the tree. Variable importance is determined by estimating how much prediction error increases when data for each variable are permuted. Random forests are considered an

**Table 1** Predictor variables for species distribution models

Predictor variable (units)	Spatial resolution (m)	Contiguous US Mean ± 95 % CI for all data	North US		South US		p for difference of means
			Mean ± 95 % CI of presences (n = 112)	Mean ± 95 % CI of pseudo-absences (n = 21,654)	Mean ± 95 % CI of presences (n = 337)	Mean ± 95 % CI of pseudo-absences (n = 7,766)	
Mean annual number of frost days <sup>a</sup>	54,000	6.5 ± 0.4	3.5 ± 4.9	7.4 ± 0.5	1.1 ± 0.8	0.28	0.0078
Mean January dew point temperature <sup>b</sup> (°C)	4,000	-5.9 ± 0.1	-7.7 ± 0.3	-8.1 ± 0.1	8.1 ± 0.5	0.38	<0.001
Annual mean precipitation <sup>b</sup> (mm)	4,000	961.6 ± 4.6	1167.4 ± 35.6	902.1 ± 5.2	1111.2 ± 9.3	<0.001	<0.001
Summer NDVI <sup>c</sup> (unitless *100)	1,000	64.0 ± 0.2	59.3 ± 3.0	66.1 ± 0.2	45.1 ± 1.3	<0.001	<0.001
Slope <sup>d</sup> (degrees)	250	1.6 ± 0.02	1.8 ± 0.1	0.5 ± 0.1	0.2 ± 0.1	<0.001	<0.001
Percent of cell covered in forest <sup>e</sup> (%)	30	33.7 ± 0.3	17.6 ± 0.4	35.5 ± 0.4	3.3 ± 0.9	<0.001	<0.001
Percent of cell covered in scrub <sup>e</sup> (%)	30	11.0 ± 0.2	1.1 ± 0.5	9.6 ± 0.2	3.0 ± 0.9	<0.001	<0.001
Percent of cell covered in crop and pasture <sup>e</sup> (%)	30	26.1 ± 0.3	5.5 ± 2.6	29.0 ± 0.4	9.8 ± 1.7	<0.001	<0.001
Housing density <sup>f</sup> (housing units/sq km)	Block groups	34.9 ± 1.4	700.2 ± 202.7	28.0 ± 1.1	221.8 ± 26.9	<0.001	<0.001
Median household income <sup>g</sup> (\$1,000)	Tracts	46.8 ± 0.1	67.4 ± 2.5	47.8 ± 0.2	50.2 ± 0.8	<0.001	<0.001
Distance to nearest railway <sup>h</sup> (km)	1,000	14.0 ± 0.2	2.7 ± 0.5	14.6 ± 0.3	9.7 ± 1.7	<0.001	<0.001
Distance to nearest town with >100,000 inhabitants <sup>i</sup> (km)	1,000	121.3 ± 1.2	20.8 ± 3.1	137.4 ± 1.6	33.5 ± 3.1	<0.001	<0.001
Distance to nearest town with >500,000 inhabitants <sup>i</sup> (km)	1,000	272.70 ± 2.16	62.6 ± 10.8	297.4 ± 2.7	286.2 ± 21.0	<0.001	<0.001
Distance to nearest interstate <sup>j</sup> (km)	1,000	39.7 ± 0.5	5.7 ± 1.1	42.8 ± 0.6	17.7 ± 2.8	<0.001	<0.001

The difference in mean values for the predictor variables between the presence and pseudo-absence points was tested using a bootstrap test with 10,000 resamples

Data source and years: <sup>a</sup> Climatic Research Unit, University of East Anglia, UK (Mitchell and Jones 2005), 1989–2009; <sup>b</sup> Prism climate group (2004), Oregon State University, US, 2007–2010; <sup>c</sup> eMODIS from USGS (Jenkerson et al. 2010), 2007–2011; <sup>d</sup> Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) from USGS (Danielson and Gesch 2011), 2011; <sup>e</sup> National Land Cover Data (Fry et al. 2011), 2006; <sup>f</sup> SILVIS Lab, University of Wisconsin-Madison, 2000; <sup>g</sup> American Community Survey (United States Census Bureau), 2010; <sup>h</sup> U.S. National Transportation Atlas Railroads, 2000; <sup>i</sup> U.S. Census (ESRI Product), 2000; <sup>j</sup> Census 2010 TIGER/Line files, 2000

improvement over single classification trees due to their high classification accuracy and their ability to handle large numbers of predictor variables, find a signal in noisy data, and identify variable importance (Cutler et al. 2007). Species distribution models were developed using randomForest version 4.6–6 in R (Liaw and Wiener 2002).

Because we were particularly interested in testing the importance of different predictor variables in different geographic regions, we analyzed three different datasets. The first dataset contained all monk parakeet and American robin occurrences in the contiguous United States. We then split the dataset geographically, along 36° N latitude, into two separate datasets. We selected 36° N latitude as our dividing line because few monk parakeets were reported in the mid-latitudes of the conterminous US (Fig. 1). The southern region included grid cells with a center latitude  $\leq 36^\circ$  N and the remainder of the grid cells were assigned to the northern region. To test for sensitivity of our results to this value, we varied the latitude at which the break occurred from 35° to 37° N. This did not substantially alter the number of monk parakeets in each subset but did affect the pseudo-absence points included in each region.

For each dataset, we ran randomForest with 5,000 trees generated for each model. Growing many trees has been shown to help improve the stability of variable importance measures (Liaw and Wiener 2002; Strobl et al. 2007). The model parameter *mtry* was set at 8, which means that 8 of our 14 variables were randomly sampled as candidates to enter the model at each split in the trees being built. Because monk parakeet prevalence is low, i.e., there are many fewer presence points compared to pseudo-absence points, we constrained the bootstrapped sample used to build each tree so that randomForest used an equal number of presence and pseudo-absence points each time. This process helped ensure that the presence and pseudo-absence points were given equal weight in the final model and thus that the model predicts monk parakeet presence with higher classification accuracy (Freeman et al. 2012). Furthermore, following the sampling approach of Wilsey et al. (2012), we randomly generated 500 datasets and built models for each using different random seeds each time. By repeating the analysis with different seeds, we were able to test the stability of the permutation-based variable importance measures (Strobl et al. 2007). For

each model run, our full sample consisted of 449 monk parakeet presence locations and an equal number of randomly selected pseudo-absences. The North and South subgroups had 112 and 337 presence points, respectively, and an equal number of pseudo-absences for each model run (Table 1).

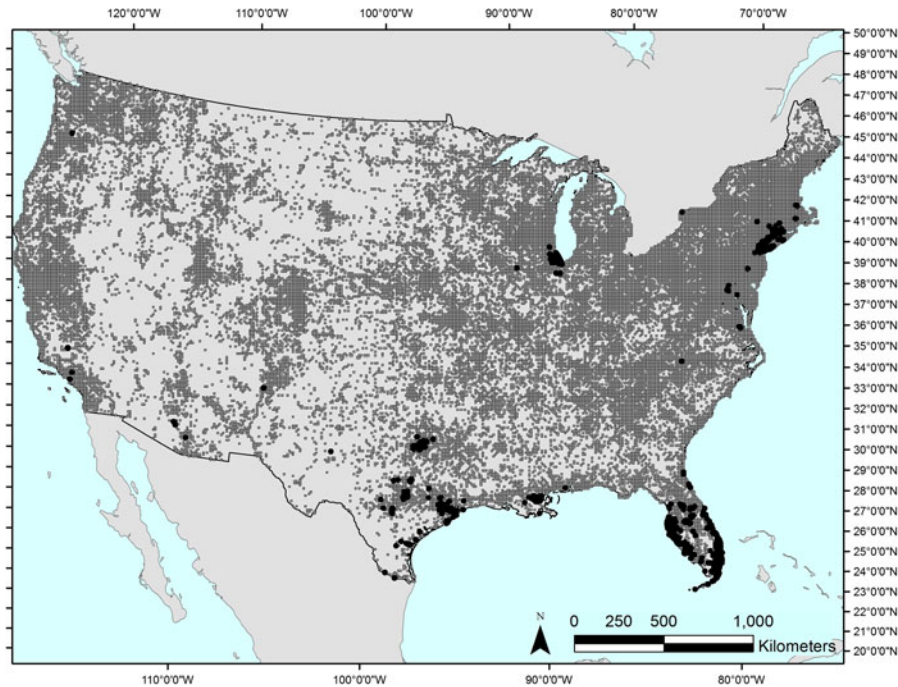
Approximately one-third of all data points were withheld from building trees and were independently run through the tree and classified. These points are referred to as the test set or “out of bag” sample. After building the 5,000 trees, each data point gets assigned to a class (presence or absence of monk parakeets in our case) based on whether that point was classified most often as a presence or an absence point when it was run through the tree as part of the test set. This majority vote becomes the “out of bag” (OOB) prediction for each observation. The OOB error estimate is then calculated by comparing the proportion of times that the majority vote is not equal to its true classification (Maidonald and Braun 2010). Another measure of model performance is Cohen’s kappa, which measures the proportion of correct predictions after accounting for chance effects (Manel et al. 2001). Kappa values from 0.6 to 0.8 indicate substantial model performance (Landis and Koch 1977 in Manel et al. 2001). Since we completed 500 randomForest runs, we report the average, minimum and maximum of both the OOB error estimates and kappa statistics (Table 2). The mean OOB error estimate and kappa statistic were used to judge the quality of the models.

Finally, the difference in mean values between the presence and pseudo-absence points for each predictor variable was tested using a bootstrap test with 10,000 resamples. All analyses were performed using the R statistical package version 2.15.0 (R Development Core Team, 2012).

## Results

There were 29,420 grid cells with American robin observations (and no monk parakeet observations), and 449 grid cells with monk parakeet observations (Fig. 1). The majority of monk parakeets can be found in Florida, Texas, New York, Illinois and Louisiana. Few monk parakeets occurred between the latitudes of 35°–37° N. By contrast American robins occurred in every state, although fewer were reported in more arid regions of the United States (Fig. 1).





**Fig. 1** Monk parakeet presence points (*black dots*,  $n = 449$ ) and American robin presences (*dark grey shading*) as reported through several citizen science datasets

### Species distribution modeling

For the full dataset, mean January dew point temperature and housing density were most important in determining presence of monk parakeets (Fig. 2c), with mean variable importance values of 0.12 and 0.06, respectively. For the northern dataset, distance to cities with more than 100,000 inhabitants and housing density were the most important variables (Fig. 2a), with mean variable importance values of 0.08 for both. For the southern dataset, mean January dew point temperature (mean variable importance = 0.15) was by far the most important variable used to differentiate between presence and pseudo-absence points (Fig. 2b). In general, human-related factors (distance to towns, interstates, and railways, household income, and housing density) were the most important factors that influenced monk parakeet distribution for the northern monk parakeets (mean importance value = 0.05), while the biophysical variables were less important (mean importance value = 0.01). Conversely, biophysical variables were more important than human-related factors (mean importance values = 0.05 and 0.02, respectively) for the southern subgroup.

Bootstrap tests were used to evaluate differences in the mean of the predictor variables between presence and pseudo-absence locations (Table 1). In the south, the mean dew point temperature in January was significantly higher ( $p < 0.001$ ) for presence points ( $\bar{x} = 8.1 \pm 0.5$  °C, 95 % CI) compared to pseudo-absences ( $\bar{x} = -0.3 \pm 0.1$  °C, 95 % CI). However, in the north, the mean January dew point temperature was not significantly different ( $p = 0.38$ ) between presences ( $\bar{x} = -7.7 \pm 0.3$  °C, 95 % CI) and pseudo-absences ( $\bar{x} = -8.1 \pm 0.1$  °C, 95 % CI). In the north, monk parakeet presences are significantly closer ( $p < 0.001$ ) to cities of more than 500,000 inhabitants ( $\bar{x} = 62.6 \pm 10.8$  km, 95 % CI) than pseudo-absences ( $\bar{x} = 297.3 \pm 2.2$  km, 95 % CI) but in the south they are significantly farther away ( $\bar{x} = 286.2 \pm 21.0$  km (95 % CI,  $p < 0.001$ ) than pseudo-absences ( $\bar{x} = 206.4 \pm 2.8$  km, 95 % CI). In both the north and the south, monk parakeets were found in areas closer to railways, interstates, and cities of more than 100,000 inhabitants, and in areas with significantly more precipitation, less summer vegetation (NDVI), less forest, less scrub, less crop and pasture, higher housing density, and higher household income (Table 1).

**Table 2** Model performance metrics

Dataset	Misclassification rate of pseudo-absences (%)	Misclassification rate of presence points (%)	Total error (%)	Kappa
ALL	8 (4.9–11.4)	6.5 (4.7–8.7)	7.2 (5.3–9.5)	0.86 (0.82–0.90)
NORTH	8.2 (2.7–14.3)	7.9 (3.6–12.5)	8.1 (4.5–12.1)	0.84 (0.75–0.91)
SOUTH	10.2 (5.9–13.9)	9 (6.5–11.6)	9.6 (6.7–12.5)	0.81 (0.76–0.87)

Mean, minimum, and maximum misclassification rates for pseudo-absence points, presence points, all points (total error rate), and kappa statistic

All the models performed well, with mean kappa values above 0.8 (Table 2), although the model of the southern United States performed slightly worse than the other two models. The mean overall OOB error rates for the full U.S., northern and southern datasets were 7.2, 8.1 and 9.6 %, respectively (Table 2). The average misclassification rates for pseudo-absences were always larger than for presences. This means that American robin grid cells were more likely to be misclassified as monk parakeet grid cells than the other way around. The error rate and variable importance did not change substantially when we varied the break between northern and southern subgroups (not shown).

## Discussion

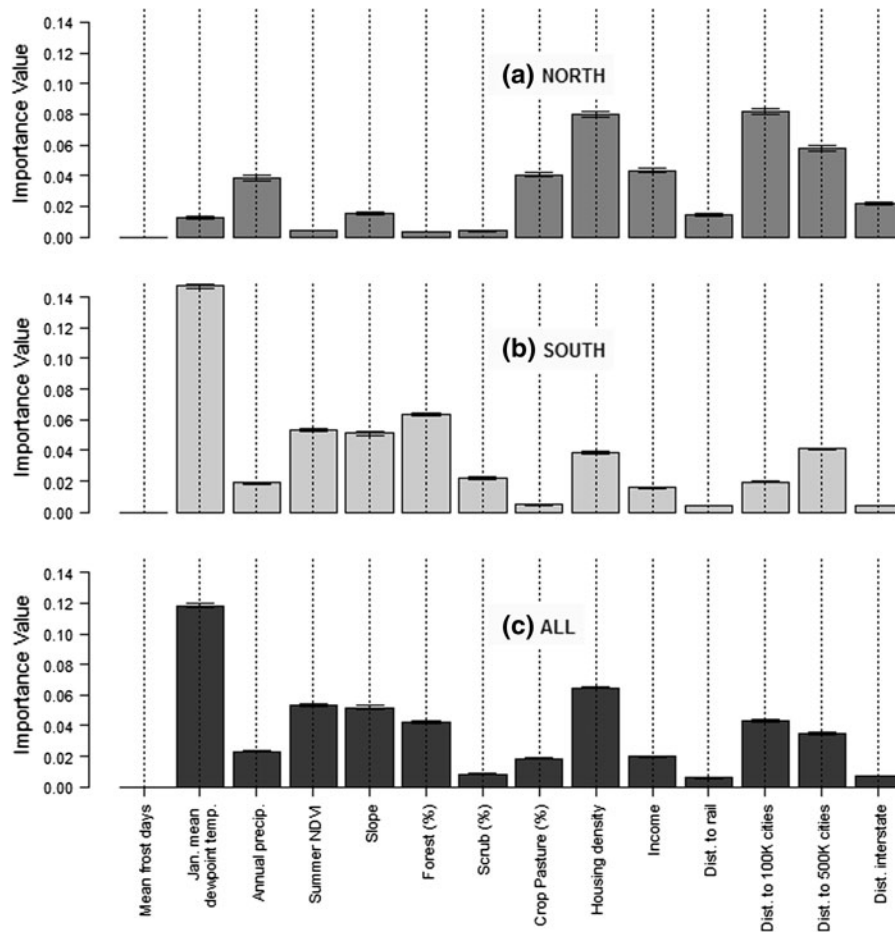
We found that a different set of variables explained the distribution of monk parakeets in the southern and northern United States (Fig. 2). Since monk parakeets were introduced through the pet trade, it is reasonable to expect that their spread would initiate from the most populated cities. Indeed, ‘distance to cities of more than 500,000 inhabitants’ and ‘housing density’ are important variables for both geographic regions. However, the southern birds have expanded beyond highly urbanized areas, as demonstrated by the fact that monk parakeet observations are significantly further away from cities of more than 500,000 inhabitants than pseudo-absences. The variables that best explain the distribution of the southern subgroup are biophysical (e.g., temperature and forest cover; Fig. 2b) rather than human-related. The southern birds are found in areas that are farther from large cities but warmer; i.e., they are found in areas that are more similar to their native range. On the other hand, distribution of the northern monk parakeets is best explained by the anthropogenic variables (Fig. 2a),

and mean January dew point temperature does not differ significantly between presence and pseudo-absence points (Table 1); i.e., the parakeets are restricted to living near cities. Had we only conducted our analyses at the scale of the contiguous United States (Fig. 2c), we would have overlooked these trends and misunderstood the factors most important to predicting monk parakeet spread.

There are two possible explanations for our finding that anthropogenic factors best explain the distribution of northern birds while biophysical factors best explain the distribution of southern birds. First, monk parakeets might have been introduced later in the northern United States and thus have not had enough time to spread beyond cities. However, we could not find any evidence that monk parakeets had different introduction dates in the north and south. Indeed, the Audubon Society’s Christmas Bird Count data suggest that the birds were first sighted in both southern states (e.g., Arizona, Florida, Louisiana, Texas) and northern states (e.g., Connecticut, Illinois, Massachusetts, New Jersey) between 1970 and 1980. This was confirmed by Lever (2006), who reports that monk parakeets were seen in northern states (Connecticut, New Jersey, Pennsylvania, and Massachusetts) in the late 1960s or early 1970s and in southern states (Florida, Alabama, and Texas) in the 1980s. We therefore reject this as a probable explanation for the importance of different explanatory variables in the two geographic regions.

A second, more likely explanation is that monk parakeets might be experiencing different pressures in different parts of the United States. According to the range map in Forshaw (2010), much of the species’ native range occurs in subtropical climates of South America. While small parts of their range extend below 40°S into the lower-elevation parts of Argentina, the climate is still relatively mild in this region (e.g., mean minimum temperature from 1961 to 1990 in Bariloche for the coldest month (July) is  $-1^{\circ}\text{C}$ ,





**Fig. 2** Variable importance measures with standard deviation *error bars* from random forest models

World Meteorological Organization). Bariloche is outside the monk parakeets' native range, in the foothills of the Andes Mountains, but is the city closest to the southern edge of their range for which we could find reliable climate information; winters are most likely colder in Bariloche than within their range. Therefore, it is surprising that populations have become established in cities in the northern United States (e.g., Chicago and Boston), where mean minimum January temperatures from 1961 to 1990 are  $-11^{\circ}\text{C}$  and  $-6^{\circ}\text{C}$ , respectively (World Meteorological Organization). It is possible that the urban heat island effect (Oke 1973) is enough to dampen the temperature variations in northern cities compared to the rural surrounding areas, making the climate more suitable for monk parakeets. Furthermore, cities offer unique and sometimes highly abundant food resources, which some species can effectively exploit.

Monk parakeets have a varied diet (South and Pruett-Jones 2000) and are known to frequent bird feeders. In Barcelona, the presence of monk parakeets is positively associated with the percentage of people over 65 years old (Rodríguez-Pastor et al. 2012), a group who may be most likely to feed birds. In colder climates, such as Chicago, birdfeeders may play an especially important role in winter. Studies of the Chicago monk parakeet population show that their diet in winter consists almost exclusively of seed from birdfeeders (South and Pruett-Jones 2000), suggesting that they rely on that resource for survival in colder climates.

It appears that the unique conditions in cities may act as “substitutable habitats” (Morellet et al. 2011; Dunning et al. 1992) for the monk parakeets, allowing them to persist in a region that might otherwise be unsuitable. Here, we extend the idea of substitutable

habitats, previously applied to a species obtaining interchangeable resources from different habitats in the same landscape (e.g., deer using hedgerows when the availability of woodland habitat decreases; Morellet et al. 2011), to exotic species using novel landscapes (e.g., cities). These novel landscapes provide resources and environmental conditions that, although very different from the species' native habitat, allow the exotic species to become established. For monk parakeets, these resources can include the food that humans provide voluntarily or involuntarily in cities but also novel nesting substrates available in urban environments. If these factors are indeed at play, we expect that monk parakeets will not expand beyond metropolitan areas in the northern United States, even though they may continue to spread and perhaps become pests in the south if they settle into agricultural areas. This has already occurred in Europe where monk parakeets have expanded from Barcelona into neighboring cropland (Domènech et al. 2003).

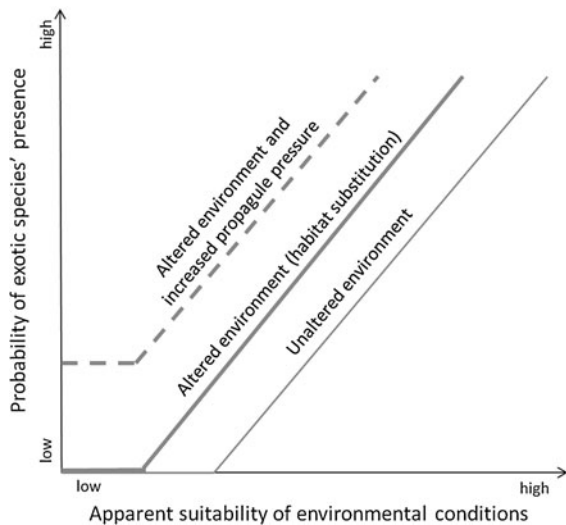
Figure 3 presents a hypothesis about factors that might allow an exotic species to survive in an apparently inhospitable region. In general, as environmental conditions become more and more similar to the species' native range, the probability of occurrence increases (thin gray line in Fig. 3). The threshold of suitability, or the conditions under which an exotic species can become established in a new locale, can shift if humans modify the environment in a way that provides substitutable habitat (thick gray line in Fig. 3). Regardless of environmental conditions, high propagule pressure may allow a species to be present although not necessarily self-sustaining (dashed gray line in Fig. 3). The specific manner in which the local environmental conditions interact with propagule pressure and human modifications of the environment is a needed future area of research, and should be examined for different exotic species. Only once we learn the weights of these factors can we most effectively manage biological invasions.

We must point out that our measures of human influence cannot distinguish between propagule pressure, supplemental resources, and altered environmental conditions. Cities can provide all of these factors (Fuller et al. 2008; Grimm et al. 2008; Chytrý et al. 2008). It is possible that the northern populations are not viable and are just replenished over time by more birds that have escaped or were released. However, the Chicago population has been studied

extensively and has been growing over the last several decades (Pruett-Jones et al. 2012), suggesting that the birds are surviving and successfully reproducing in the region. Several surrogates that approximate propagule pressure have been proposed but unfortunately most of them also affect establishment stages (Chiron et al. 2009), so that identifying the mechanisms by which humans affect exotic species richness remains challenging (Richardson and Pyšek 2008). Future research on exotic monk parakeets could mark and recapture individual birds (following the marking methods of Senar et al. 2012) and monitor nests (particularly in northern populations) to estimate reproduction, mortality, and immigration rates. If available, other explanatory variables might further our understanding of human influence on monk parakeet establishment. In particular, locations of urban parks, sales of bird seed, or sales of monk parakeet pets could be essential in understanding the interplay between the pressures of finding suitable and available habitat, obtaining proper resources for survival, and propagule pressure on the establishment of this particular exotic species.

An assumption of our analysis is that monk parakeets have had an opportunity to colonize our entire study area. We acknowledge that this is unlikely to be the case, as birds are less likely to escape or be released in uninhabited or sparsely populated parts of the country. In particular, some areas in the western United States (e.g., parts of the Great Basin and Rocky Mountain regions) have few human residents and also few reports of monk parakeets (Fig. 1). However, these same regions had few American robin observations as well, and thus carried little weight in our model. Because our analysis focused on a relative comparison of variable importance between the northern and southern regions of the United States, and because our approach excluded remote locations where humans do not live or visit, we believe that this assumption should have minimal impact on our findings.

Other studies have reported the importance of human influences in the range expansion of monk parakeets (Muñoz and Real 2006; Strubbe and Matthysen 2009) but our study goes one step further and disentangles factors that will be important in predicting the future spread of this species. In particular, we show that while the distribution of northern monk parakeets may be limited to urban environments, the distribution of the southern subgroup is best explained by biophysical variables. This research highlights the



**Fig. 3** Conceptual diagram of mechanisms responsible for exotic species presence in different environments

importance of understanding the regionally-specific factors that allow an exotic species to become established, which is key to predicting their expansion beyond areas of introduction.

**Acknowledgments** This research was supported by the National Science Foundation and the U.S. Forest Service under the ULTRA-Ex program (NSF Grant Number 0948484). We thank members of the Minor Lab group for providing feedback on earlier drafts of the manuscript.

## References

- Arnfield AJ (2003) Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. *Int J Climatol* 23:1–26
- Beyer HL (2004) Hawth's analysis tools for ArcGIS. <http://www.spatial ecology.com/htools>. Accessed 27 Dec 2012
- Blackburn TM, Duncan RP (2001) Determinants of establishment success in introduced birds. *Nature* 414(6860): 195–197
- Bomford M, Kraus F, Barry SC et al (2009) Predicting establishment success for alien reptiles and amphibians: a role for climate matching. *Biol Invasions* 11:713–724
- Bonter DN, Cooper CB (2012) Data validation in citizen science: a case study from Project FeederWatch. *Front Ecol Environ* 10:305–307
- Breiman L (2001) Random forests. *Mach Learn* 45:5–32
- Butler CJ (2003) Population biology of the introduced Rose-ringed Parakeet *Psittacula krameri* in the UK. Department of Zoology, University of Oxford, UK, p 332
- Catford JA, Jansson R, Nilsson C (2009) Reducing redundancy in invasion ecology by integrating hypotheses into a single theoretical framework. *Divers Distrib* 15:22–40
- Chiron F, Shirley S, Kark S (2009) Human-related processes drive the richness of exotic birds in Europe. *P R Soc B* 276:47–53
- Chytrý M, Jarosik V, Pyšek P et al (2008) Separating habitat invisibility by alien plants from the actual level of invasion. *Ecology* 89(6):1541–1553
- Colautti RI, Grigorovich IA, MacIsaac HJ (2006) Propagule pressure: a null model for biological invasions. *Biol Invasions* 8:1023–1037
- Cutler DR, Edwards TC, Beard KH et al (2007) Random forests for classification in ecology. *Ecology* 88(11):2783–2792
- Danielson JJ, Gesch DB (2011) Global multi-resolution terrain elevation data 2010 (GMTED2010): U.S. Geological Survey Open-File Report 2011–1073
- Davis LR (1974) The monk parakeet: a potential threat to agriculture. In: Proceedings of the 6th vertebrate pest conference (1974). Paper 7. <http://digitalcommons.unl.edu/vpc6/7>
- Dickinson JL, Zuckerberg B, Bonter DN (2010) Citizen science as an ecological research tool: challenges and benefits. *Annu Rev Ecol Evol Syst* 41:149–172
- Domènech J, Carrillo-Ortiz J, Senar JC (2003) Population size of the monk parakeet *Myiopsitta monachus* in Catalonia. *Revista Catalana d'Ornitologia* 20:1–9
- Dormann CF, Elith J, Bacher S (2012) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*. doi:10.1111/j.1600-0587.2012.07348.x
- Dunning JB, Danielson BJ, Pulliam HR (1992) Ecological processes that affect populations in complex landscapes. *Oikos* 65:169–175
- Forshaw JM (2010) Parrots of the world. Princeton University Press, Princeton, NJ
- Freeman EA, Moisen GG, Frescino TS (2012) Evaluating effectiveness of down-sampling for stratified designs and unbalanced prevalence in random forest models of tree species distributions in Nevada. *Ecol Model* 233:1–10
- Fry J, Xian G, Jin S, Dewitz J, Homer C, Yang L, Barnes C, Herold N, Wickham J (2011) Completion of the 2006 national land cover database for the conterminous United States. *Photogramm Eng Rem S77(9)*:858–864
- Fuller RA, Warren PH, Armsworth PR et al (2008) Garden bird feeding predicts the structure of urban avian assemblages. *Divers Distrib* 14:131–137
- Grimm NB, Faeth SH, Golubiewski NE et al (2008) Global change and the ecology of cities. *Science* 319:756–760
- Hulme PE (2009) Trade, transport and trouble: managing invasive species pathways in an era of globalization. *J Appl Ecol* 46:10–18
- Hyman J, Pruett-Jones S (1995) Natural history of the monk parakeet in Hyde Park, Chicago. *Wilson Bull* 107(3): 510–517
- Jenkerson CB, Maiersperger T, Schmidt G (2010) eMODIS: a user-friendly data source: US Geological Survey Open-File Report 2010–1055
- Jeschke JM, Strayer DL (2006) Invasion success of vertebrates in Europe and North America. *P Natl Acad Sci USA* 102(20):7198–7202
- Lepczyk CA, Mertug AG, Liu J (2004) Assessing land owner activities related to birds across rural-to-urban landscapes. *Environ Manage* 33:110–125

- Leprieur F, Beauchard O, Blanchet O et al (2008) Fish invasions in the world's river systems: when natural processes are blurred by human activities. *PLoS Biol* 6:404–410
- Lever C (2006) *Naturalized birds of the world*. Longman Scientific and Technical, New York
- Levine J, Adler P, Yelenik S (2004) A meta-analysis of biotic resistance to exotic plant invasions. *Ecol Lett* 7:975–989
- Liaw A, Wiener M (2002) Classification and regression by random forest. *R News* 2(3):18–22
- MacLeod CJ, Newson SE, Blackwell G, Duncan RP (2009) Enhanced niche opportunities: can they explain the success of New Zealand's introduced bird species? *Divers Distrib* 15:41–49
- Maindonald J, Braun WJ (2010) *Data Analysis and Graphics Using R: An Example-Based Approach*, 3rd edn. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge, Massachusetts, USA
- Manel S, CeriWillimans H, Ormerod SJ (2001) Evaluating presence-absence models in ecology: the need to account for prevalence. *J Appl Ecol* 38:921–931
- Mateo RG, Croat TB, Felicísimo AM, Muñoz J (2010) Profile or group discriminative techniques? Generating reliable species distribution models using pseudo-absences and target-group absences from natural history collections. *Divers Distrib* 16(1):84–94
- McKinney M (2006) Urbanization as a major cause of biotic homogenization. *Biol Conserv* 127(3):247–260
- Minor ES, Appelt CW, Grabiner S, Ward L, Moreno A, Pruet Jones S (2012) Distribution of exotic monk parakeets across an urban landscape. *Urban Ecosyst* 15:979–991
- Mitchell TD, Jones PD (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int J Climatol* 25:693–712
- Møller AP (2009) Successful city dwellers: a comparative study of the ecological characteristics of urban birds in the Western Palearctic. *Oecologia* 159:849–858
- Morellet N, Van Moorter B, Cargnelutti B et al (2011) Landscape composition influences roe deer habitat selection at both home range and landscape scales. *Landscape Ecol* 26:999–1010
- Moyle PB, Marchetti MP (2006) Predicting invasion success: freshwater fishes in California as a model. *Bioscience* 56:515–524
- Muñoz AR, Real R (2006) Assessing the potential range expansion of the exotic monk parakeet in Spain. *Divers Distrib* 12(6):656–665
- Munson MA, Caruana R, Fink D, Hochachka WM, Iliff M, Rosenberg KV, Sheldon D, Sullivan BL, Wood C, Kelling C (2010) A method for measuring the relative information content of data from different monitoring protocols. *Methods Ecol Evol* 1:263–273
- Murray JV, Stokes KE, Klinken RD (2012) Predicting the potential distribution of a riparian invasive plant: the effects of changing climate, flood regimes and land-use patterns. *Glob Chang Biol* 18:1738–1753
- Oke TR (1973) City size and the urban heat island. *Atmos Environ* 7:769–779
- Peterson AT (2003) Predicting the geography of species' invasions via ecological niche modeling. *Q Rev Biol* 78:419–433
- Phillips SJ, Dudík M, Elith J, Graham C, Lehmann A, Leathwick J, Ferrier S (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecol Appl* 19:181–197
- PRISM Climate Group, Oregon State University (2004). <http://prism.oregonstate.edu>. Accessed 27 Dec 2012
- Pruett-Jones S, Appelt CW, Sarfaty A et al (2012) Urban parakeets in Northern Illinois: a 40-year perspective. *Urban Ecosyst* 15(3):709–719
- Pyšek P, Jarosik V, Hulme PE et al (2010) Disentangling the role of environmental and human pressures on biological invasions across Europe. *P Natl Acad Sci USA* 107(27):12157–12162
- Rebele F (1994) Urban ecology and special features of urban ecosystems. *Glob Ecol Biogeogr* 4(6):173–187
- Richardson DM, Pyšek P (2008) Fifty years of invasion ecology: the legacy of Charles Elton. *Divers Distrib* 14:161–168
- Rodríguez-Pastor R, Senar JC, Ortega A et al (2012) Distribution patterns of invasive monk parakeets (*Myiopsitta monachus*) in an urban habitat. *Anim Biodivers Conserv* 35(1):107–117
- Roura-Pascual N, Hui C, Ikeda T et al (2011) Relative roles of climatic suitability and anthropogenic influence in determining the pattern of spread in a global invader. *P Natl Acad Sci USA* 108(1):220–225
- Russello MA, Avery ML, Wright TF (2008) Genetic evidence links invasive monk parakeet populations in the United States to the international pet trade. *EvolBiol* 8:217
- Senar JC, Carrillo-Ortiz J, Arroyo L (2012) Numbered neck collars for long-distance identification of parakeets. *J Field Ornithol* 83:180–185
- Sexton JP, McIntyre PJ, Angert AL et al (2009) Evolution and ecology of species range limits. *Annu Rev Ecol Evol S* 40:415–436
- Shea K, Chesson P (2002) Community ecology theory as a framework for biological invasions. *Trends Ecol Evol* 17:170–176
- Simberloff D (2009) The role of propagule pressure in biological invasions. *Annu Rev Ecol Evol Syst* 40:81–102
- Sol D, Santos DM, Fera E, Clavell J (1997) Habitat selection by the monk parakeet *Myiopsitta monachus* during the colonization of a new area. *Condor* 99:39–46
- Sol D, Bartomeus I, Griffin AS (2012) The paradox of invasion in birds: competitive superiority or ecological opportunism? *Oecologia* 169(2):553–564
- Song U, Mun S, Ho CH, Lee EJ (2012) Responses of two invasive plants under various microclimate conditions in the Seoul metropolitan region. *Environ Manage* 49:1238–1246
- South JM, Pruet Jones S (2000) Patterns of flock size, diet, and vigilance of naturalized monk parakeets in Hyde Park, Chicago. *Condor* 102(4):848–854
- Spreyer MF, Bucher EH (1998) Monk parakeet (*Myiopsitta monachus*). In: Poole A, Gill F (eds) *The birds of North America*, No. 322. The Birds of North America, Inc., Philadelphia, PA
- Strobl C, Boulesteix AL, Zeileis A, Hothorn T (2007) Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics* 8:25–46
- Strubbe D, Matthysen E (2009) Establishment success of invasive ring-necked and monk parakeets in Europe. *J Biogeogr* 36(12):2264–2278
- Sukopp H (2004) Human-caused impact on preserved vegetation. *Landsc Urban Plan* 68:347–355

- Sullivan BL, Wood CL, Iliff MJ, Bonney RE, Fink D, Kelling S (2009) eBird: a citizen-based bird observation network in the biological sciences. *Biol Conserv* 142: 2282–2292
- Taylor BW, Irwin RE (2004) Linking economic activities to the distribution of exotic plants. *Proc Natl Acad Sci USA* 101(51):17725–17730
- Thuiller W, Richardson DM, Pyšek P et al (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Glob Chang Biol* 11:2234–2250
- Thuiller W, Richardson DM, Rouget M et al (2006) Interactions between environment, species traits, and human uses describe patterns of plant invasions. *Ecology* 87:1755–1769
- Tilman D (2004) A stochastic theory of resource competition, community assembly and invasions. *Proc Natl Acad Sci USA* 101:10854–10861
- Van Bael S, Pruett-Jones S (1996) Exponential population growth of monk parakeets in the United States. *Wilson Bull* 108:584–588
- vanHeezik Y, Smyth A, Mathieu R (2008) Diversity of native and exotic birds across an urban gradient in a New Zealand city. *Landsc Urban Plan* 87:223–232
- Wilsey CB, Lawler JJ, Cimprich DA (2012) Performance of habitat suitability models for the endangered black-capped vireo built with remotely-sensed data. *Remote Sens Environ* 119:35–42
- World Meteorological Association (n.d.). World Weather Information Service, Retrieved May 30, 2013 from <http://worldweather.wmo.int/>
- Yaukey PH (2010) Citizen science and bird-distribution data: an opportunity for geographical research. *Geogr Rev* 100(2): 263–273