

Using citizen-reported data to predict distributions of two non-native insect species in Sweden

LINA A. WIDENFALK,¹ KARIN AHRNÉ,² AND ÅSA BERGGREN^{1,†}

¹Swedish University of Agricultural Sciences, Department of Ecology, Box 7044, SE-75007 Uppsala, Sweden

²Swedish University of Agricultural Sciences, Swedish Species Information Centre, Box 7007, SE-75007 Uppsala, Sweden

Citation: Widenfalk, L. A., K. Ahrné and A. Berggren. 2014. Using citizen-reported data to predict distributions of two non-native insect species in Sweden. *Ecosphere* 5(12):156. <http://dx.doi.org/10.1890/ES14-00212.1>

Abstract. Information on exotic species' current and potential distribution is vital for decisions on management. Species distribution models can predict where colonizations are likely; however, the collection of species' distribution data over large areas in order to parameterize the models is costly. Therefore, modelling methods that are able to use low-cost information such as citizen-reported data are potentially very useful. In this study, we used the species habitat modelling program Maxent to predict the potential geographical distribution of two non-native insect species in Sweden, the butterfly *Araschnia levana* and the shield bug *Graphosoma lineatum*. For this we used citizen-reported presence-only open-access data in combination with climate and land cover data from national databases. Our models showed that presence of *A. levana* was best predicted by winter temperature and habitats related to open grasslands. For *G. lineatum*, summer temperature and open green areas, in both urban and rural areas were the best predictors for species presence. These models show that large areas of non-colonized potential habitats exist within Sweden. For *A. levana* these yet-to-be-colonized habitats are mainly in the south, while for *G. lineatum* these habitats occur in the south and along the Baltic Sea coast. Comparisons of temporal patterns in species reporting for *A. levana* and *G. lineatum* to similar insects with known stable populations revealed large 'willingness to report' effects that could potentially bias range expansion rates. Once corrected for, current distribution expansion rates were estimated as 1.9 km/yr and 1.07 km/yr respectively. The study shows the use of public reports in conservation science as a way of gathering species information over large areas. This increases the data sources available for researchers to predict the distribution of species and have the additional value of the involvement of the public in conservation efforts.

Key words: *Araschnia levana*; citizen science data; distribution model; exotic species; *Graphosoma lineatum*; grid occupancy; surveillance; Sweden.

Received 4 July 2014; revised 16 October 2014; accepted 17 October 2014; final version received 20 November 2014; **published** 22 December 2014. Corresponding Editor: D. P. C. Peters.

Copyright: © 2014 Widenfalk et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. <http://creativecommons.org/licenses/by/3.0/>

† **E-mail:** asa.berggren@slu.se

INTRODUCTION

Predicting and quantifying current and potential distributions of exotic species is a critical step in evaluating their environmental impacts and management control options (Beale et al. 2008, Drury and Rothlisberger 2008, Keller et al. 2008). To date, this task has been generally achieved by

examining species presence/absence data in order to determine the ecological factors that limit or permit current and future distributions (Guisan and Thuiller 2005, Elith and Leathwick 2009). However, collecting high quality data over large distribution areas is costly; thus, most studies are restricted to using presence-only or extent-of-occurrence data, which is the minimum

convex polygon drawn to encompass all the known, inferred or projected sites of occurrence, including cases of vagrancy (IUCN 2001). For species whose distributions encompass large areas or are rapidly changing, species-specific monitoring programs may be cost-prohibitive. Because of the growing need for distribution data but a lack of resources to gather it, newer methods have been developed that can utilize presence-only data collected by the general public and reported to species databases (Gormley et al. 2011).

Citizen-reported species data from many countries are available in open source databases (e.g., <http://www.pwrc.usgs.gov/bbs/>, <http://invasives.biodiversityireland.ie>, <http://data.nbn.org.uk>). These data often represent an enormous collective effort over large regions and are potentially valuable sources of information for studies on species ecology and distribution (Aslan and Rejmánek 2010, Kéry et al. 2010), but are under-used in research and management (Goffredo et al. 2010). Despite the potential usefulness of citizen-collected data (Cacho et al. 2010, Cacho and Hester 2011) they are generally not gathered through structured sampling methods and may be heavily influenced by observers' willingness to report (Snäll et al. 2011). This may bias reporting, give misleading impressions of species distribution shifts as reporting trends change and limit the types of questions that can be answered. In addition, citizen-reported data generally lack absence records, which are needed for most species distribution models (Elith et al. 2006). Thus, presence-only data methods have been developed (e.g., the software Maxent; Phillips et al. 2006) to largely overcome this problem. Maxent is considered to be one of the best performing methods for modelling species distributions compared with species distribution models (GARP; Stockwell and Peters 1999), generalized linear models (McCullagh and Nelder 1989), boosted regression trees (Friedman et al. 2000) and random forests (Breiman 2001, as cited in Elith and Graham 2009). The method has also been successfully used to model invasive species distributions (Elith et al. 2006, 2010, Wang et al. 2010, Gormley et al. 2011) and citizen-reported data offer a potentially huge (although largely untested) resource for modelling species distributions and range expansions.

The Swedish Species Information Centre (Art-Databanken) operates the Swedish Species Observations System (www.artportalen.se; hereafter referred to as 'the species database'): a species database to which the general public, organizations and authorities can report species observations throughout the country. For example, in September 2013 the database for terrestrial and limnetic invertebrates had records of over 2.6 million observations. Data reported to the species database are checked by national species' experts (taxonomists) working at or affiliated with the center, and exceptional reports are individually validated. Because Sweden is at the northern range of many species' distributions and the country is separated from mainland Europe by the Baltic Sea, Sweden is a good place to examine range expansions of native species associated with current climate change and recently introduced species (Betzholtz et al. 2013). We used data from the species database to predict species distributions for insects currently expanding their range, and specifically examined the potential distribution of two species that are non-native to Sweden: the map butterfly *Araschnia levana* L. and the striped shield bug *Graphosoma lineatum* L. Currently, the invasive potential of these species is unknown.

We used Maxent to analyze citizen-reported species presence data together with climate and land cover data from national databases to examine the species-environment relationships, with the aim to increase our knowledge of their present and potential future distributions in Sweden. We included climate factors in our analyses in addition to land-cover variables because at high latitudes insects are often limited in their distributions by low temperatures affecting the survival of adults and dormant life-history stages (Tran et al. 2007). Thus our general aim was to predict current and future distributions in Sweden of two non-native insect species with expanding ranges, with these analyses a demonstration of the potential value of citizen-reported data for such predictions. Specifically we were interested in: (1) which climate and land cover data are currently correlated with the presence of *A. levana* and *G. lineatum* in Sweden, and (2) based on the habitat and climate correlations with current distributions, where are the likely areas available for colonization by

these two species in Sweden? In addition, we used the temporal and geographical changes in reporting of these insects to the species database to examine whether citizen-reported data could provide direct empirical support for range expansion, and if the current expansion rate could be estimated. This relied on estimating and adjusting for potential reporting bias (i.e., increased willingness to report observations to the species database in recent years, independent of range expansion) by comparing the reporting trends for *A. levana* and *G. lineatum* that are assumed to be undergoing range expansion, to reporting trends for similar insects known to have a stable distribution in Sweden.

MATERIAL AND METHODS

Study species

Araschnia levana (the map butterfly; family Nymphalidae) is found in central and northern Europe, and northern Asia (Eliasson et al. 2005). In Scandinavia it is currently established in Denmark, Sweden and Finland (Vliegenthart et al. 2011); in Finland it was first recorded in 1973 and has been increasing its distribution rapidly northwards since the 1980s (Mitikka et al. 2008). The species was first recorded in Sweden 1982 and has had a rapid northward expansion here as well; it is now one of the most common butterfly species in southern Sweden (Eliasson et al. 2005). The butterfly over-winters as pupae in Sweden and it generally has two or three generations per year. The first generation is on the wing during May to June, the second generation is flying from July to August, and if a third generation emerges it is flying in September. The map butterfly is often found in small openings in forest areas (V. Mitikka and B. Söderström, *personal communication*) and along small roads and small clear-cuts where the host plant, *Urtica dioica* L. (stinging nettle), is abundant (Eliasson et al. 2005). The butterfly's summer generation is more numerous than the spring generation and is believed to be the primary dispersal phase (Fric and Konvicka 2000, 2002). *Araschnia levana* is considered a medium to good disperser (Bink 1992), and a positive relationship between annual dispersal and temperature in late summer has been found in Finland (Mitikka et al. 2008), indicating that climate may affect the expansion rate of the

species. Previous attempts to predict range shifts in the species using only bioclimatic models performed poorly, indicating that other environmental variables may also be of importance (Mitikka et al. 2008).

Graphosoma lineatum is a phytophagous shield bug species (family Pentatomidae) distributed from central Asia to southern and central Europe and northern Africa. It is considered a Mediterranean species, but with a northward expanding distribution (Bringmann 1977). In Scandinavia the bug is currently established in Denmark, Sweden and Finland (Aukema and Rieger 2006, Berend 2011) and since its establishment in 1916 *G. lineatum* has increased its distribution in southern Sweden (C.-C. Coulianos, *personal communication*). It hibernates in the ground and when emerging in mid-May, the coloration is strikingly red and black striped (Gamberale-Stille et al. 2010). Reproduction occurs in June and July and the nymphs develop in early July to August (Tullberg et al. 2008). In Sweden *G. lineatum* is found on sun-exposed meadows of *Anthriscus sylvestris* L. (wild chervil) and *Aegopodium podagraria* L. (ground elder), but also on other garden plants within the *Apiaceae* family (Tullberg et al. 2008). The striped shield bug does not have a particular dispersal period; it is believed to move during the whole season to locate suitable forage areas (Nakamura 1998). There is however a difference in motion level between pre- and post-hibernating adults. In the post-hibernation stage bugs are more active, especially males searching for mates, while bugs in the pre-hibernation stage are cryptic, focused on feeding and are sedentary (Johansen et al. 2011). Previous experimental studies show that flight activity is positively correlated to day-length but that the species has no migratory behavior (Nakamura 1998). Thus, *G. lineatum* is thought to move only minimum distances needed to find mates, new feeding-plants and hibernation shelters (Nakamura 1998, Tillman et al. 2009). Their gregarious habits also make them less likely to benefit from moving long distances (Vesely et al. 2006).

Modelling method and evaluation

We used Maxent 3.3.3 k (Phillips et al. 2006, Phillips and Dudík 2008) to model the potential geographic distributions of *A. levana* and *G. lineatum*. Maxent is a machine learning method

that fits a probability distribution for species occurrence to pixels or sites in a specified geographic area – this allows accurate estimates of species geographic distributions to be derived using presence-only data (Elith et al. 2006, 2011). It finds the probability distribution of maximum entropy, i.e., the probability distribution that is most spread out or closest to uniform given the environmental constraints derived from the occurrence data (Phillips et al. 2006, Phillips and Dudík 2008). The software uses presence-only data and a pre-defined number of randomly selected points (pseudo-absences) in combination with environmental covariates to construct an index of habitat suitability for each cell ranging from 0 (least suitable) to 1 (most suitable habitat; see Phillips et al. 2006, Phillips and Dudík 2008, Elith et al. 2011 for more details).

We used the default settings in Maxent (Phillips and Dudík 2008) unless otherwise specified. We used 50% of the data for the random tests and replicated the runs ten times. This form of replication repeatedly splits the data into random training and testing subsets. We chose maximum test sensitivity and specificity as threshold rule to optimize the sum of model sensitivity and specificity and to reduce the rate of false negatives and positives (Manel et al. 2001, Hernandez et al. 2006). We created response curves to investigate how the predicted relative probability of occurrence depends on the value of each environmental variable. Furthermore, we used the jack-knife option to measure the importance of all variables used, by doing the analyses with each environmental variable first omitted then used in isolation. To assess the performance of the Maxent model we plotted a receiver operating characteristic curve, which compares the model's sensitivity (true positives) against 1 – specificity (false positives) over the entire range of thresholds. When modelling with presence-only data the area under this curve (AUC) represents the probability that a randomly chosen presence site will be ranked more suitable than a randomly chosen pseudo-absence site. Although AUC has its weaknesses (Lobo et al. 2008), it is the predominantly used method to measure the performance of predictive distribution models (Elith and Graham 2009). A model that performs no better than random will have an AUC of 0.5, whereas a model with perfect

discrimination will have an AUC of 1 (Fielding and Bell 1997).

Species data

The species data were obtained from the Swedish Species Observations System (www.artportalen.se). The Species Observations System is a database provided by the Swedish Species Information Centre (ArtDatabanken), to which the general public (that heavily dominate the reporting), organizations and authorities can report species observations. The observations also include data on geographical position, abundance and in some cases data on life-history stage. The reported observations are checked by taxonomic specialists working at the center or affiliated taxonomists. Observations that are positioned outside the species' current distribution, reported at an abnormal time of the season or have some other unusual qualities are individually validated. A secondary type of validation often occurs when a species is reported outside its distribution, as other reporters driven by their own curiosity go to the site and check the individual/individuals and report back to the database. At the time of the study, the database for invertebrates had approximately 2.6 million observations reported. We used all species data with unique coordinates for the two study species. For *A. levana* the first observation in the species database was from 1985 (with a gap of 10 years to the next observation) and the first observation for *G. lineatum* was from 1963. As climate data were only available up to 2008, we used species observation data between the first year of observation for each species (using 1995 as the first year for *A. levana*) to year 2008 in the models. This resulted in a total of 1103 observations for *A. levana* and 685 observations for *G. lineatum* (Fig. 1). The distributions of the reported data for both species were verified by taxonomists (B. Söderström and C.-C. Coulianos, *personal communication*).

Environmental predictors

Land cover data.—For both insect species, we included land cover data from the whole of Sweden. The data were obtained from the SMD (Swedish land cover data) database, which is the refined Swedish national version of the CORINE land cover database (Engberg 2005). The SMD

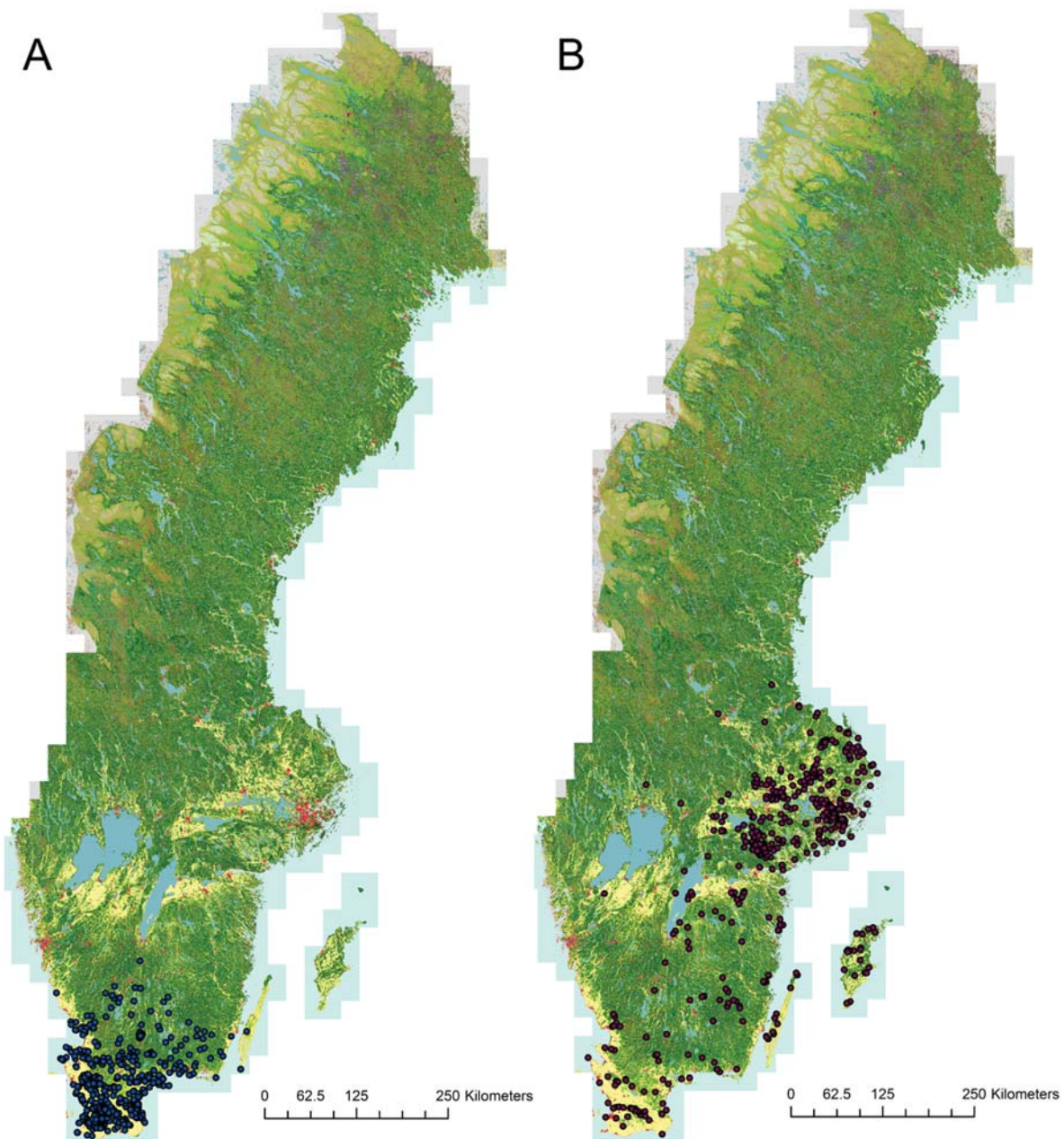


Fig. 1. The species presence locations used to build the Maxent models to predict the species' potential geographic distributions. The occurrences of (A) *A. levana* (1995–2008) and (B) *G. lineatum* (1963–2008) reported to the database the Swedish Species Observations System. Background map: ©Lantmäteriet, ärende nr I 2010/0345.

gives information on land use, land cover and vegetation and is a homogeneous database with about 60 land cover classes. All land cover classes except open water were used in the analyses. The smallest mapping unit in this set is 1–25 ha (consistent within each category) and the map

resolution is 25×25 m.

Climate data.—We used climate data for the whole of Sweden in the analyses. The variables were chosen based on previous studies and expert knowledge of the two species. When modelling the potential distribution of *A. levana*

we included four climate variables: mean July temperature, mean August temperature, mean winter temperature (i.e., December to February) and mean annual water balance. These variables have all been suggested as important in determining the distribution of the map butterfly (Mitikka et al. 2008) and were calculated as mean yearly values for the period 1995 to 2008. When modelling the distribution of *G. lineatum* we used two climate variables: the mean August temperature and mean September temperatures; these variables were chosen as proxy for the number of sun hours during late summer, which may be important for adult development and survival during winter (Slachta et al. 2002; C.-C. Coulianos, *personal communication*). Previous studies have shown a high tolerance towards low temperatures, so winter temperature was not included in the models for *G. lineatum* (Hodkova and Hodek 2004). Data on temperature, annual precipitation and evaporation for the years 1963–2008 were obtained from the Swedish Meteorological and Hydrological Institute (SMHI). For *A. levana* we used climate data for the years 1995–2008, and for *G. lineatum* we used climate data from 1963 to 2008. The temperature data came as gridded data sets with 4 km between points. As all environmental layers used in Maxent must have the same geographic bounds and cell size, we used the Interpolate tool in the toolbox ‘Spatial Analyst Tools’ in ArcMap 9.3 (Environmental Systems Research Institute, Redlands, California, USA) to interpolate the data to a raster map with cell size 25 × 25 m. The mean annual water balance was calculated as annual precipitation subtracted by annual evaporation. The data on precipitation and evaporation was given for each run-off area (total 1052 run-off areas) in Sweden and came as a map in vector format with each run-off areas as a separate polygon. The map was converted from polygon to raster, with a cell size of 25 × 25 m, using the Conversion Tools tool set in Arc Map 9.3. The raster maps of mean temperatures and water balance were then converted from raster to ASCII format, using the same tool set.

Estimating species distribution expansion

Although range expansion of the two focal species is strongly suggested by increased reports in recent years to the national species database

(Eliasson et al. 2005; C.-C. Coulianos, *personal communication*), it is unclear how much of this reflects reporting bias, because citizen-reported data may be poorly correlated with actual presence when the willingness to report a species changes with time (Snäll et al. 2011). Expansion rate is a potentially useful measure as it is related to the speed with which new areas will be colonized in the future; thus, it is important to examine how citizen-reported data can be used to estimate these rates and whether reporting biases can be adjusted for. To do this we first extracted positional and temporal data from the species database for the two study species. This gave us yearly data on the number of observations and the area of colonization being reported, as all observations were mapped to 1 × 1 km squares covering the distribution area. Temporal changes in the reported distribution area can be used to calculate the expansion rate per year at the range margin (Hill et al. 2001, Preuss et al. 2014), which is derived from the slope of a function fitted to the area occupied per year (i.e., marginal velocity = slope/√pi). However, such a calculation relies on the assumption that increases in the reported distribution area reflect true changes in distribution, rather than an increased willingness to report observations. Thus, we used records from three species related to each of the study species to account for any reporting bias of the study species: for *A. levana* we used the butterflies *Papilio machaon* L., *Argynnis paphia* L. and *Maculinea arion* (L.); for *G. lineatum* we used the shield bugs *Dolycoris baccarum* (L.), *Pentatoma rufipes* (L.) and *Eurydema oleracea* (L.). These species were chosen as being the best candidates for detecting reporting bias because they: (1) were related to the study species and were similarly attractive and recognizable; thus should have been equally attractive for reporting to the database, (2) had long reporting histories and were frequently reported, and (3) had stable population distributions, and thus could be used as a reference for reporting bias. Data on these were retrieved from the taxonomists working with the Species Observations System and the national Red List at the Swedish Species Information Centre. Species distribution, threat status and ecology are continuously evaluated at the center (B. Söderström and J. Sandström, *personal communication*).

For the range expansion estimates for the two focal species we fitted a sigmoidal function predicting the square root of the area of occupancy per year, in the form:

$$V_{\max} \times \frac{\text{year}^n}{h^n + \text{year}^n}$$

where V_{\max} is the value where the function asymptotes, h is the point on the x-axis (in this case the year) which corresponds to the maximal rate of increase, and n is the rate of increase. From this function we could calculate the marginal range expansion (km/yr) for each year by finding the first-order derivative for the function (i.e., the slope of the line) and dividing this by $\sqrt{\pi}$ (Hill et al. 2001). To account for reporting bias, we down-adjusted the slope based on the difference between slopes calculated from sigmoidal functions fitted to the number of observations reported per year for the focal species, and the mean of the corresponding ‘control’ species (see previous section). For example, if there was no difference in the slopes of the functions fitted to observations of the focal species and the control species for a particular year, then this suggests any range expansion is completely confounded by willingness-to-report bias in that year (with the expansion rate being related to the difference between the two slopes). For all range expansion analyses we fitted functions, their first-order derivatives (i.e., calculated slopes) and derived predictions using an MCMC Gibb’s sampler (JAGS; Plummer 2003) called from R (R Core Team 2014). Thus, estimates are Bayesian posterior distributions; these allow statements to be made regarding the probability of differences between quantities of interest, and allow 95% CIs to be easily calculated for these quantities. In all cases 50000 iterations were run, with multiple chains and a burn-in of 10000; convergence was checked by visual inspection of chain plots and the Gelman and Rubin diagnostic (Gelman and Rubin 1992).

RESULTS

Habitat-climate associations and predicted distributions

Maxent performed well in the species-environment relationships for both *A. levana* and *G. lineatum*. The AUC scores, demonstrating useful-

ness of the models, were relatively high for both species: 0.881 ± 0.007 for *A. levana* and 0.886 ± 0.004 for *G. lineatum*. According to the model predictions there are many suitable regions for both species within the country that are not yet occupied. In the very south of Sweden and areas northwards along the coasts there are unoccupied habitat for *A. levana* (Fig. 2B), while for *G. lineatum* the model showed suitable habitat for the species across the entire southern part of Sweden as well as along the full length of the East Coast (Fig. 3B).

The probability of presence of *A. levana* decreased as winter temperature lowered. At a mean winter temperature below -1.7°C the probability of presence was zero. The variables contributing most to the Maxent model for *A. levana* were winter temperature and land cover data (Table 1). In jack-knife tests the variable with highest gain when used in isolation was winter temperature (AUC = 0.821 ± 0.010), while land cover data was the variable that decreased the gain most when omitted (AUC = 0.839 ± 0.008) (see Table 2 for AUC values of all the compared models). The land cover classes most strongly related to presence of *A. levana* were: non-urban parks, moors and heathland, gravel and sand pits, shrubland, and pastures (see Table 3 for description of the classes). They can all be considered as important for the species as the probability of presence of *A. levana* was consistently (for the 10 replicated runs) high, i.e., above 0.6. High values were sometimes obtained for other land cover classes as well, but showed large variation between runs and were therefore considered less important.

The probability of presence of *G. lineatum* increased when temperatures in August and September increased. At mean August temperatures below 11.5°C or mean September temperatures below 5.5°C the probability of presence of *G. lineatum* was zero. The variables contributing most to the model for *G. lineatum* were land cover data and August temperature. Because there is a correlation between August and September temperature, the ordering of the two should be interpreted with caution (Table 1). In jackknife tests land cover was the variable with the highest gain when used in isolation (AUC = 0.788 ± 0.008) and it was the variable that decreased the gain most when omitted (AUC = 0.835 ± 0.008)

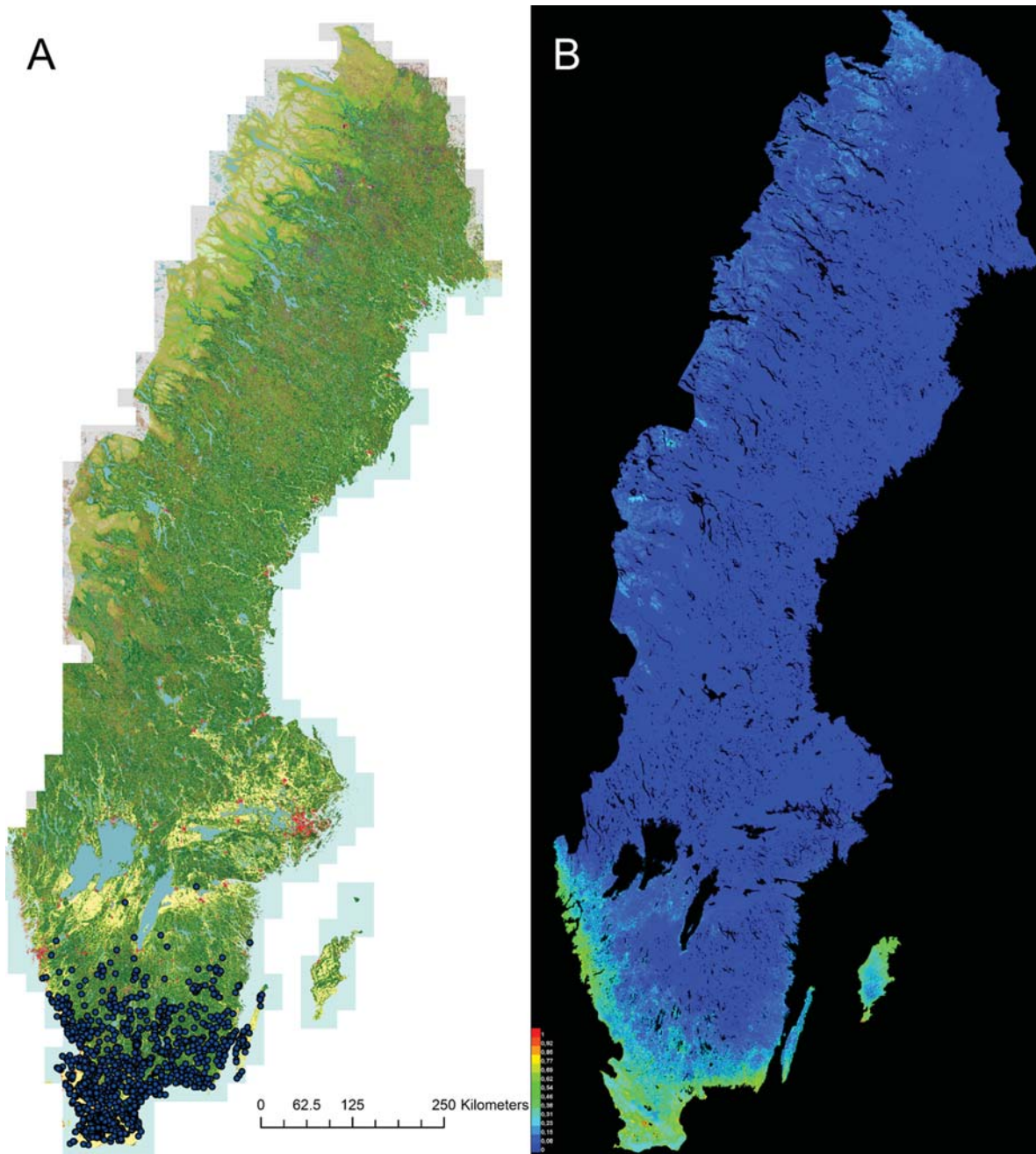


Fig. 2. Comparison of (A) the current distribution of *A. levana*, based on the observations reported in the Swedish Species Observations System 1985–2013 (Background map: ©Lantmäteriet, ärende nr I 2010/0345), with (B) the predicted maximum potential distribution of *A. levana* in Sweden when analyzed with Maxent (scale ranging from blue = very unsuitable to red = very suitable).

(Table 2). The land cover classes most strongly related to the presence of *G. lineatum* were: non-urban parks, small holdings with land, gravel

and sand pits, urban green areas, pastures and urban areas with large green areas (see Table 3 for description of the classes). In all these six land

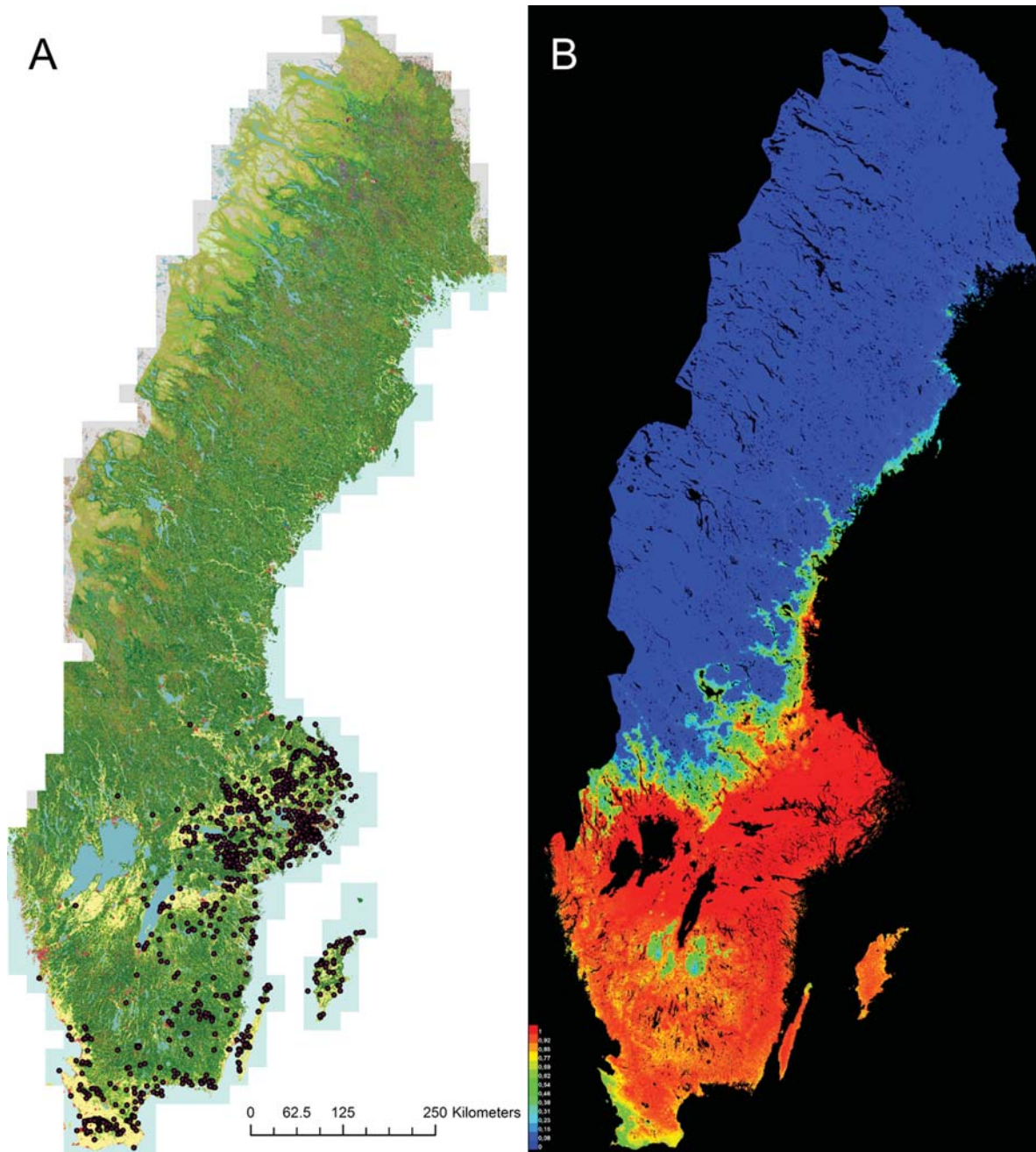


Fig. 3. Comparison of (A) the current distribution of *G. lineatum*, based on the observations reported in the Swedish Species Observations System 1963–2013 (Background map: ©Lantmäteriet, ärende nr I 2010/0345), with (B) the predicted maximum potential distribution of *G. lineatum* in Sweden when analyzed with Maxent (scale ranging from blue = very unsuitable to red = very suitable).

cover classes the probability of presence of *G. lineatum* was consistently (for the 10 replicated runs) high, above 0.6.

Distribution expansion and reporting bias

There was strong evidence of reporting bias from changes in ‘willingness-to-report’ observa-

Table 1. Percent contribution to the final Maxent models by all the different variables used for each of the two species.

Variable	Contribution (%)
<i>Araschnia levana</i>	
Mean winter temperature 1995–2008	53.4
Land cover data	36.4
Mean water balance 1995–2008	7.4
Mean temperature July 1995–2008	1.6
Mean temperature August 1995–2008	1.2
<i>Graphosoma lineatum</i>	
Land cover data	45.3
Mean temperature August 1963–2008	34.4
Mean temperature September 1963–2008	20.3

Table 2. The resulting AUC values (mean \pm SE) when each variable is excluded in turn and the model is created with the remaining variables (column “AUC without variable”) and when each variable is used in isolation to create the model (column “AUC with variable only”).

Variable	AUC without variable	AUC with variable only
<i>Araschnia levana</i>		
Mean winter temperature 1995–2008	0.868 \pm 0.008	0.821 \pm 0.010
Land cover data	0.839 \pm 0.008	0.797 \pm 0.010
Mean water balance 1995–2008	0.878 \pm 0.007	0.728 \pm 0.017
Mean temperature July 1995–2008	0.880 \pm 0.007	0.723 \pm 0.011
Mean temperature August 1995–2008	0.879 \pm 0.007	0.770 \pm 0.012
<i>Graphosoma lineatum</i>		
Land cover data	0.835 \pm 0.008	0.788 \pm 0.008
Mean temperature August 1963–2008	0.839 \pm 0.005	0.725 \pm 0.008
Mean temperature September 1963–2008	0.848 \pm 0.004	0.719 \pm 0.009

tions to the Species Observations System for the control species with stable populations (Fig. 4). Reported observations were only occasional until 2003 for butterflies and 2005 for bugs; for the following 5 or so years reports followed an exponential growth pattern before reaching an asymptote at around 2008–2010. There was a corresponding exponential growth in reports of the focal non-native species at a similar time (Fig. 4), suggesting that range expansion during much of the period in this study is severely confounded by reporting bias caused by changes in willingness to report. However during the last years of the study, the plateauing of reported observations of control species with a corresponding continued growth of the non-native species observations provides strong evidence that the focal species are expanding their distributions. Based on this, we can cautiously estimate the rate of range expansion for the focal species based on the last two years of the study, as willingness-to-report-bias approaches zero (i.e., the slope of the control species model approaches zero). Range expansion rate estimates (as a mean for 2009 and

2010) are: 1.9 km/yr for the map butterfly (95% CI 0.79–3.31) and 1.07 km/yr for the shield bug (95% CI 0.34–2.03).

DISCUSSION

Predicted distributions and environment-habitat associations

Climate data was an important predictor of the distribution for both species (Table 1). For *A. levana* winter temperature was the most important predictor of species presence, while for *G. lineatum* land cover was the most important. It appears that low winter temperatures make it difficult for over-wintering stages of *A. levana* to survive and establish in northern and colder latitudes (Mitikka et al. 2008). Because Sweden has on average a 2°C lower mean temperature at similar latitudes to Finland (Tveito et al. 2000), this would explain why the distribution of some insects show a higher northern distribution in Finland and the neighboring Russia, than in Sweden (Vanhanen et al. 2007, Mitikka et al. 2008).

Table 3. The land cover classes most strongly related to the presence of the two species, with description and SMD-code.

Land cover class	Code	Description
Non-urban parks	1.4.2.5	Parks outside urban areas such as parks surrounding estates
Moors and heathland	3.2.2	Areas with low vegetation, dominated by low shrubs, bushes and herbs
Small holdings with land	1.1.2.3	Areas with single or groups of houses on the countryside, where 30–80% of the area is artificial surfaces
Gravel and sand pits	1.3.1.1	Pits with adjoining buildings and roads, only naturally occurring sand and gravel
Urban green areas	1.4.1	Urban areas with more than 70% cover of vegetation, such as urban parks, cemeteries, allotments and botanical gardens
Shrubland	3.2.4.1	Areas with bushes 1–5 m high with >30% cover
Pastures	2.3.1	Grassland used for grazing or mowing, with less than 30% trees and bushes
Urban areas with larger green areas	1.1.2.1.2	Areas where 50–70% of the land consists of green areas, mainly private gardens

For the shield bug the most important predictor was land cover, particularly open green areas, which is consistent with the species' ecology (Tullberg et al. 2008). It can therefore be expected that such areas near the range margin for *G. lineatum* will be colonized in coming years. For *A. levana*, urban grasslands, shrubland and pastures, moors and heathlands, and naturally occurring sand and gravel areas are important

habitat areas, with both urban grasslands and pastures often containing the larval food plant, stinging nettle (Eliasson et al. 2005, Mitikka et al. 2008). Shrubland, sand and gravel areas are habitats that can be found both in open and forested areas and the importance of these land cover classes for the map butterfly could be explained by its preference for forest glades and areas influenced by human activity. The impor-

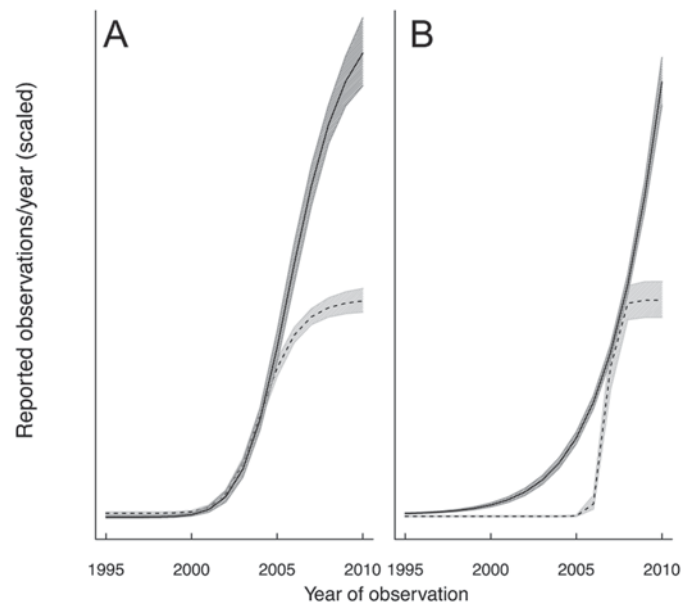


Fig. 4. Observations per year reported to the Swedish Species Observations System between 1995 and 2010 for (A) *A. levana* (black line) and comparative butterfly species with stable populations (dashed line), and (B) *G. lineatum* (black line) and comparative bug species with stable populations (dashed line). Lines represent modeled predictions from a sigmoid function (see text for details) and shaded areas are 95% CIs. The y-axis has been rescaled for ease of comparison with the final values being 420 and 306 for *A. levana* and other butterflies, respectively, and 759 and 64 for *G. lineatum* and other bugs, respectively.

tance of moors and heathlands, which are normally associated with dry and nutrient poor areas, is not as obvious; however, the SMD land cover class also includes moist areas and the importance of this land cover class may simply be explained by the preference of *A. levana* for open areas. The model predicts that the suitable areas for the butterfly mainly exist in the southern part of the country. While for the shield bug, suitable areas for colonization exist across southern Sweden and along the east coast (Figs. 2 and 3). However, predictions of suitable areas for these species that are currently expanding their distribution, may be less certain than the models suggests because the magnitude of habitat preference is underestimated due to false pseudo-absences (i.e., the species have not had time to colonize all preferred areas; Václavík and Meentemeyer 2009). If so, it could be that *A. levana* has not had time to colonize areas with lower temperatures, and would explain why we see *A. levana* being reported in some areas after 2008 that were predicted as unsuitable by the model (Fig. 2). Previous studies with invasive species have found that Maxent is well-suited to predict habitat associations and that spatial and detection biases in the presence-only data are low (Gormley et al. 2011), because the general species-environmental relationships remain the same. Knowing this, the potential underestimation of suitable colonization areas should be kept in mind when evaluating model results of a spreading species. We expect that for *A. levana* and *G. lineatum* we will see an on-going expansion into habitats identified in our models. Because the species we studied are sensitive to temperature, we expect expansion of their distributions to be primarily in the southern (warmer) parts of the country, where populations are still to establish in much of their preferred habitat. In the north however, distribution changes will be largely related to local temperatures, with the harsher climate making species expansion towards colder latitudes slow and the edge of the distributions fluctuate between years from variations in temperature (Berggren et al. 2009). With an estimated increase in mean temperature in regions at higher latitudes in the coming decades (Solomon et al. 2007), we may expect a higher likelihood of the species being established in the regions that are not suitable

today (Jiménez-Valverde et al. 2011).

Distribution expansion and reporting bias

Estimates of yearly range expansion are a function of two key components, each of which may influence the distribution expansion estimate: i.e., the yearly measure calculated from observational data and the model fitted to these data to estimate the temporal trend (Preuss et al. 2014). One of the major problems with citizen-reported data is that interest in a species and the willingness to report it may change over time (Snäll et al. 2011), and hence make it seem that a species is expanding its distribution (if reports are increasing) independent of any changes in the species' true distribution (Jeppsson et al. 2010). Thus it is critical that these biases are acknowledged and, if possible, accounted for when using citizen-collected data for estimating species range expansion. In this study we took advantage of the fact that the Species Observations System database contained information on many species that had stable population distributions to estimate reporting bias, and from this we could either: (1) adjust our estimates based on the reporting bias calculated from species with stable populations, and or (2) find periods where the reporting of 'stable' species had plateaued and look for evidence of range expansion in the focal species during that time. Although there was strong evidence that range expansion data for the focal species were confounded by reporting biases (Fig. 4), there was also strong evidence for range expansion as the patterns of reporting stable species and the focal species deviated in recent years (i.e., stable species reports levelling-off, while the focal species reports continue to climb). These consistent patterns for each of the three control species for both groups (bugs and butterflies), suggests that such patterns may be easily detectable and, thus, controlled for in range expansion estimation using citizen-collected data.

For the map butterfly our range expansion estimate was 0.79–3.31 km/yr, which is very similar to that reported in Finland from survey data from 1983 to 1998 (1.29 and 1.47 km/yr), but less than that reported since 1999 (7.5 km/yr; Mitikka et al. 2008). Like in Finland prior to 1998, the Swedish population may still be in the early population growth phase, with this potentially

going to increase in coming years. However, one should also consider that our estimate is too low because sampling of the occupied range is likely to be less than complete, and expansion rates based on grid occupancy decline with incomplete sampling (e.g., Melles et al. 2011, Preuss et al. 2014). Thus, both the butterfly and shield bug range expansion estimates should be considered as minimums, with the true expansion rates likely to be higher than this. Although studies on phytophagous shield bugs similar to *G. lineatum* have shown that individuals move only short distances (mainly to the next food patch) a yearly range expansion estimate of 1.07 km is not surprising because of general differences between individual dispersal rates and species expansion rates. Species expansion is the sum of the movement dynamics of geographically distributed populations where extinction and colonization events are controlled both by species' ecology and the environment, and may not be directly predictable by individual movement patterns (Berggren et al. 2009, Preuss et al. 2011, Kaňuch et al. 2013).

If we can generalize our results to species of similar size, distribution ability and life-histories (Peacock and Worner 2008) we would expect a similar occupancy pattern and increasing distribution of butterflies and shield bugs non-native to the region. When attempting to manage non-native species, one of the main aims is to understand possible future distributions (where can it go?) and the time it will take for the species to establish in new areas (when will it get there?). By using a combination of public data with the predictive function in Maxent and the temporal geographical distribution of the species, we believe we have reasonably estimated these two factors. However, calculating rate of change of the distribution is potentially more difficult. One reason is that any predictions are made based on earlier distribution of the species. Previous distribution is not necessarily a good predictor of future change as it, e.g., may not represent well the species niche (Jiménez-Valverde et al. 2011), and this uncertainty increases with length of time predicted. Therefore, even though the distribution increase of the study species is based on large amounts of species position data over long time periods, forecasting adds additional uncertainties.

In addition to increasing our knowledge of the habitat preferences of two non-native species, our study shows that citizen-reported presence-only data can be extremely valuable when predicting distributions and estimating whether species are undergoing range shifts. This study could not have been carried out without the large amount of voluntarily collected data, which presented an opportunity to work on large spatial and temporal scales that would not be available for most research projects. Future utilization of such data resources will aid ecological understanding and allow new insights on species ecology in relatively short time frames (Aslan and Rejmánek 2010, Kelly et al. 2012); however, these data are far from perfect and need to be analyzed with an understanding of their deficiencies. By using the public to report new species, governments may be able to more efficiently allocate resources for managing invasive organisms (Cacho and Hester 2011). We encourage other researchers and managers to analyze these types of data for several reasons. First it is available for many species and often covers both long time periods and large geographical areas (Eyre et al. 2004, Silvertown 2009, Meehan et al. 2013). Second, it represents information where governmental funding or public efforts have already been invested, and so represents an efficient use of resources. Third, the more these data are examined, the better we will come to understand the types of questions best suited to these data and how to best overcome any deficiencies. Finally, the use of data collected by citizens gives the contributors a sense of satisfaction and ownership of the science being conducted (Cohn 2008, Chapin et al. 2011). This, in turn, aids environmental education and increases the general knowledge of current ecological and environmental questions (Goffredo et al. 2010). Ultimately this may create positive feedbacks within the community that increases our knowledge of species and their ecology as citizens receive more encouragement to observe and submit their observations to such databases.

ACKNOWLEDGMENTS

We are grateful to all reporters to the Swedish Species Observations System (Artportalen). We would like to thank Bo Söderström, Jonas Sandström, Carl-Cedric Coulianos, Varpu Mitikka and Christer Sol-

breck that contributed with their knowledge about the study species. We also thank Matt Low for useful comments on the language and help with the range expansion modelling.

LITERATURE CITED

- Aslan, C. E., and M. Rejmánek. 2010. Avian use of introduced plants: ornithologist records illuminate interspecific associations and research needs. *Ecological Applications* 20:1005–1020.
- Aukema, B., and C. Rieger, editors. 2006. Catalogue of the heteroptera of the palaeartic region. Volume 5. Pentatomomorpha II. The Netherlands Entomological Society, Ponsen & Looijen, Wageningen, The Netherlands.
- Beale, R., J. Fairbrother, A. Inglis, and D. Trebeck. 2008. One biosecurity: a working partnership. Quarantine and Biosecurity Review Panel, Department of Agriculture, Fisheries and Forestry, Australia. <http://apo.org.au/node/2926>
- Berend, A. 2011. Fauna Europea: *Graphosoma lineatum*. Fauna Europea. Version 2.6.2. http://www.fauaaur.org/full_results.php?id=454007
- Berggren, Å., C. Björkman, H. Bylund, and M. P. Ayres. 2009. The distribution and abundance of animal populations in a climate of uncertainty. *Oikos* 118:1121–1126.
- Betzholtz, P.-E., L. B. Pettersson, N. Ryrholm, and M. Franzén. 2013. With that diet you will go far: trait-based analysis reveals a link between rapid range expansion and nitrogen-favoured diet. *Proceeding of the Royal Society B* 280:20122305.
- Bink, F. A. 1992. *Ecologische atlas van de dagvlinders van Noordwest-Europa*. Schuyt and Co Uitgevers en Importeurs, Haarlem, the Netherlands.
- Bringmann, H. 1977. Faunistic notice part 39 *Graphosoma-lineatum* (Heteroptera Pentatomidae) in northern areas of East Germany. *Entomologische Nachrichten* 21:175.
- Cacho, O. J., and S. M. Hester. 2011. Deriving efficient frontiers for effort allocation in the management of invasive species. *Agricultural and Resource Economics* 55:72–89.
- Cacho, O. J., D. Spring, S. Hester, and R. Mac Nally. 2010. Allocating surveillance effort in the management of invasive species: a spatially-explicit model. *Environmental Modelling & Software* 25:444–454.
- Chapin, F. S., M. E. Power, S. T. A. Pickett, A. Freitag, J. A. Reynolds, R. B. Jackson, D. M. Lodge, C. Duke, S. L. Collins, A. G. Power, and A. Bartuska. 2011. Earth stewardship: science for action to sustain the human-earth system. *Ecosphere* 2:89.
- Cohn, J. P. 2008. Citizen science: can volunteers do real research? *BioScience* 58:192–197.
- Drury, K. L. S., and J. D. Rothlisberger. 2008. Offense and defense in landscape-level invasion control. *Oikos* 117:182–190.
- Eliasson, C. U., N. Ryrholm, M. Holmer, K. Jilg, and U. Gärdenfors. 2005. Nationalnyckeln till Sveriges flora och fauna. Fjärilar. Dagfjärilar. Hesperidae–Nymphalidae. Artdatabanken, SLU, Uppsala, Sweden.
- Elith, J., and C. H. Graham. 2009. Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography* 32:66–77.
- Elith, J., and J. R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual review of Ecology, Evolution and Systematics* 40:677–697.
- Elith, J., M. Kearney, and S. Phillips. 2010. The art of modeling range-shifting species. *Methods in Ecology and Evolution* 1:330–342.
- Elith, J., S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of Maxent for ecologists. *Diversity and Distributions* 17:43–57.
- Elith, J., et al. 2006. Novel methods improve predictions of species' distributions from occurrence data. *Ecography* 29:129–151.
- Engberg, A. 2005. Product specification of the Swedish CORINE land cover data [Produktspecifikation av Svenska CORINE marktäckedata]. SCMD-0001, issue 2.1. the Land Survey [Lantmäteriet], Gävle, Sweden.
- Eyre, M. D., S. P. Rushton, M. L. Luff, and M. G. Telfer. 2004. Predicting the distribution of ground beetle species (Coleoptera, Carabidae) in Britain using land cover variables. *Journal of Environmental Management* 72:163–174.
- Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24:38–49.
- Fric, Z., and M. Konvicka. 2000. Adult population structure and behaviour of two seasonal generations of the European map butterfly, *Araschnia levana*, species with seasonal polyphenism (Nymphalidae). *Nota Lepidopterologica* 23:2–25.
- Fric, Z., and M. Konvicka. 2002. Generations of the polyphenic butterfly *Araschnia levana* differ in body design. *Evolutionary Ecology Research* 4:1017–1032.
- Friedman, J. H., T. Hastie, and R. J. Tibishirani. 2000. Additive logistic regression: A statistical view of boosting. *Annals of Statistics* 28:337–374.
- Gamberale-Stille, G., A. I. Johansen, and B. S. Tullberg. 2010. Change in protective coloration in the striated shieldbug *Graphosoma lineatum* (Heteroptera: Pentatomidae): predator avoidance and generalization among different life stages. *Evolutionary Ecology* 24:423–432.
- Gelman, A., and D. B. Rubin. 1992. Inference from

- iterative simulation using multiple sequences. *Statistical Science* 7:457–472.
- Goffredo, S., F. Pensa, F. Neri, A. Orlandi, M. Scola Gagliardi, A. Velardi, C. Piccinetti, and F. Zaccanti. 2010. Unite research with what citizens do for fun: “recreational monitoring” of marine biodiversity. *Ecological Applications* 20:2170–2187.
- Gormley, A. M., D. M. Forsyth, P. Griffioen, M. Lindeman, D. S. L. Ramsey, M. P. L. Scroggie, and L. Woodford. 2011. Using presence-only and presence-absence data to estimate the current and potential distributions of established invasive species. *Journal of Applied Ecology* 48:25–34.
- Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8:993–1009.
- Hill, J. K., Y. C. Collingham, C. D. Thomas, D. S. Blakely, D. M. Fox, and B. Huntley. 2001. Impacts of landscape structure on butterfly range expansion. *Ecology Letters* 4:313–321.
- Hernandez, P. A., C. H. Graham, L. L. Master, and D. L. Albert. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* 29:773–785.
- Hodkova, M., and I. Hodek. 2004. Photoperiod, diapause and cold-hardiness. *European Journal of Entomology* 101:445–458.
- IUCN. 2001. IUCN red list categories and criteria: version 3.1. IUCN Species Survival Commission. IUCN, Gland, Switzerland and Cambridge, UK.
- Jeppsson, T., A. Linde, U. Gärdenfors, and P. Forslund. 2010. The use of historical collections to estimate population trends: A case study using Swedish longhorn beetles (Coleoptera: Cerambycidae). *Biological Conservation* 143:1940–1950.
- Jiménez-Valverde, A., A. Decae, and M. A. Arnedo. 2011. Environmental suitability of new reported localities of the funnelweb spider *Macrothele calpeiana*: an assessment using potential distribution modeling with presence-only techniques. *Journal of Biogeography* 38:1213–1223.
- Johansen, A. I., B. S. Tullberg, and G. Gamberale-Stille. 2011. Motion level in *Graphosoma lineatum* coincides with ontogenetic change in defensive colouration. *Entomologia Experimentalis et Applicata* 141:163–167.
- Kañuch, P., Å. Berggren, and A. Cassel-Lundhagen. 2013. Colonisation history of *Metrioptera roeselii* in northern Europe indicates human-mediated dispersal. *Journal of Biogeography* 40:977–987.
- Keller, R. P., K. Frang, and D. M. Lodge. 2008. Preventing the spread of invasive species: economic benefits of intervention guided by ecological predictions. *Conservation Biology* 22:80–88.
- Kelly, J. F., J. R. Shipley, P. B. Chilson, K. W. Howard, W. F. Frick, and T. H. Kunz. 2012. Quantifying animal phenology in the aerosphere at a continental scale using NEXRAD weather radars. *Ecosphere* 3:16.
- Kéry, M., B. Gardner, and C. Monnerat. 2010. Predicting species distributions from checklist data using site-occupancy models. *Journal of Biogeography* 37:1851–1862.
- Lobo, J. M., A. Jiminéz-Valverde, and R. Real. 2008. AUC: misleading measure of the performance of predictive models. *Global Ecology and Biogeography* 17:145–151.
- Manel, S., H. C. Williams, and S. J. Ormerod. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology* 38:921–931.
- McCullagh, P., and J. A. Nelder. 1989. Generalized linear models. Chapman and Hall, London, UK.
- Meehan, T. D., J. Glassberg, and C. Gratton. 2013. Butterfly community structure and landscape composition in agricultural landscapes of the central United States. *Journal of Insect Conservation* 17:411–419.
- Melles, S. J., M. J. Fortin, K. Lindsay, and D. Badzinski. 2011. Expanding northward: influence of climate change, forest connectivity, and population processes on a threatened species’ range shift. *Global Change Biology* 17:17–31.
- Mitikka, V., R. K. Heikkinen, M. Luoto, M. B. Araujo, K. Saarinen, J. Pöyry, and S. Fronzek. 2008. Predicting range expansion of the map butterfly in Northern Europe using bioclimatic models. *Biodiversity and Conservation* 17:623–641.
- Nakamura, K. 1998. Effect of photoperiod on flight activity in *Graphosoma lineatum* (Heteroptera: Pentatomidae). *European Journal of Entomology* 95:297–300.
- Peacock, L., and S. P. Worner. 2008. Biological and ecological traits that assists establishment of alien invasive insects. *New Zealand Plant Protection* 61:1–7.
- Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190:231–259.
- Phillips, S. J., and M. Dudík. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31:161–175.
- Plummer, M. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In K. Hornik, F. Leisch, and A. Zeileis, editors. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, March 20–22, 2003, Vienna, Austria. <http://www.ci.tuwien.ac.at/Conferences/DSC-2003/Proceedings/>
- Preuss, S., A. Cassel-Lundhagen, and Å. Berggren. 2011. Modelling the distribution of the invasive Roesel’s bush-cricket (*Metrioptera roeselii*) in a

- fragmented landscape. *NeoBiota* 11:33–49.
- Preuss, S., M. Low, A. Cassel-Lundhagen, and Å. Berggren. 2014. Evaluating range-expansion models for calculating nonnative species' expansion rate. *Ecology and Evolution* 4:2812–2822.
- R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Silvertown, J. 2009. A new dawn for citizen science. *Trends in Ecology and Evolution* 24:467–471.
- Slachta, M., J. Vambera, H. Zahradnicková, and V. Kostál. 2002. Entering diapause is a prerequisite for successful cold-acclimation in adult *Graphosoma lineatum* (Heteroptera: Pentatomidae). *Journal of Insect Physiology* 48:1031–1039.
- Snäll, T., O. Kindvall, J. Nilsson, and T. Pärt. 2011. Evaluating citizen-based presence data for bird monitoring. *Biological Conservation* 144:804–810.
- Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, editors. 2007. Contribution of working group I to the forth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, UK.
- Stockwell, D., and D. Peters. 1999. The GARP modeling system: problems and solutions to automated spatial prediction. *International Journal of Geographic Information Sciences* 13:143–158.
- Tillman, P. G., T. D. Northfield, R. F. Mizell, and T. C. Riddle. 2009. Spatiotemporal patterns and dispersal of stink bugs (Heteroptera: Pentatomidae) in peanut-cotton farmscapes. *Environmental Entomology* 38:1038–1052.
- Tran, J. K., T. Ylioja, R. F. Billings, J. Régnière, and M. P. Ayres. 2007. Impact of minimum winter temperatures on the population dynamics of *Dendroctonus frontalis*. *Ecological Applications* 17:882–899.
- Tullberg, B. S., G. Gamberale-Stille, T. Bohlin, and S. Merilaita. 2008. Seasonal ontogenetic colour plasticity in the adult striated shield bug *Graphosoma lineatum* (Heteroptera) and its effect on detectability. *Behavioral Ecology and Sociobiology* 62:1389–1396.
- Tveito, O. E., E. Førland, R. Heino, I. Hanssen-Bauer, H. Alexandersson, B. Dahlström, A. Drebs, C. Kern-Hanssen, T. Jónsson, E. Vaarby Laursen, and Y. Westman. 2000. Nordic temperature maps. DNMI Report 09/00 KLIMA. Norwegian Meteorological Institute, Oslo, Norway.
- Václavík, T., and R. K. Meentemeyer. 2009. Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distribution? *Ecological Modelling* 220:3248–3258.
- Vanhanen, H., T. O. Veteli, S. Päivinen, S. Kellomäki, and P. Niemelä. 2007. Climate change and range shifts in two insect defoliators: gypsy moth and nun moth: a model study. *Silva Fennica* 41:621–638.
- Veselý, P., S. Veselá, R. Fuchs, and J. Zrzavý. 2006. Are gregarious red-black shieldbugs, *Graphosoma lineatum* (Hemiptera: Pentatomidae), really aposematic? An experimental approach. *Evolutionary Ecology Research* 8:881–890.
- Vliegthart, A., R. Verovnik, and M. Wiemers. 2011. Fauna Europaea: *Araschnia levana*. In O. Karsholt and E. J. van Nieukerken, editors. Fauna Europaea: Lepidoptera, Papilionoidea. Version 2.6.2. http://www.faunaeur.org/full_results.php?id=441672
- Wang, X. Y., X. L. Huang, L. Y. Jiang, and G. X. Qiao. 2010. Predicting potential distribution of chestnut phylloxerid (Hemiptera: Phylloxeridae) based on GARP and Maxent ecological niche models. *Journal of Applied Entomology* 134:45–54.