



# Where to search: the use of opportunistic data for the detection of an invasive forest pest

Maarten de Groot · Nikica Ogris ·  
Mark van der Meij · Michael J. O. Pocock

Received: 2 June 2021 / Accepted: 26 June 2022 / Published online: 16 July 2022  
© The Author(s) 2022

**Abstract** Early detection is important for the management of invasive alien species. In the last decade citizen science has become an important source of such data. Here, we used opportunistic records from the “LIFE ARTEMIS” citizen science project, in which people submitted records from places where they observed tree pests, to understand the distribution of a rapidly-spreading forest pest: the oak lace bug (*Corythucha arcuata*) in Slovenia. These citizen science records were not distributed randomly. We constructed a species distribution model for *C. arcuata* that accounted for the biased distribution of citizen science by using the records of other tree pests and diseases from the same project as pseudo-absences (so-called constrained pseudo-absences), and compared this to a model with pseudo-absences selected randomly from across Slovenia. We found that the constrained pseudo-absence model showed that *C. arcuata* was more likely to be found in east, in places with more oak trees and at lower elevations,

and also closer to highways and railways, indicating introduction and dispersal by accidental human transport. The outputs from the model with random pseudo-absences were broadly similar, although estimates from this model tended to be higher and less precise, and some factors that were significant (proximity to minor roads and human settlements) were artefacts of recorder bias, showing the importance of taking the distribution of recording into account wherever possible. The finding that *C. arcuata* is more likely to be found near highways allows us to design advice for where future citizen science should be directed for efficient early detection.

**Keywords** Oak lace bug · Ecological modelling · Citizen science · Invasive alien species · Early detection · *Corythucha arcuata*

## Introduction

Invasive alien species are a threat to biological diversity and ecosystems (Vila et al. 2011). The number of alien species is still increasing and does not appear to have reached saturation globally (Seebens et al. 2017). Although not all alien species will become invasive, 15% are regarded as invasive (Hulme 2009). Invasive alien species affect biodiversity and ecosystem services through competition, predation, crop damage, hybridisation or as disease vectors (Lowe et al. 2000). It is estimated that introduction of

---

M. de Groot (✉) · N. Ogris  
Department of Forest Protection, Slovenian Forestry  
Institute, Večna pot 2, 1000 Ljubljana, Slovenia  
e-mail: maarten.degroot@gozdis.si

M. van der Meij  
Aeres University of Applied Sciences, Arboretum West 98,  
1325 WB Almere, The Netherlands

M. J. O. Pocock  
UK Centre for Ecology & Hydrology,  
Wallingford OX10 8BB, Oxfordshire, UK

invasive alien pests can negatively impact up to 10% of the European total carbon stores (Seidl et al. 2018), substantially reduce growth rate and yields of crops (Battisti et al. 2014; Marcolin et al. 2021) and many species have been driven extinct because of invasive alien species (Clavero and García-Berthou 2005), especially on islands (Spatz et al. 2017). It is therefore important to deal rapidly with invasive alien species and an important first step is to provide effective early detection.

Early detection of invasive alien species is important to reduce their chance of establishing or spreading by facilitating rapid control and eradication measures. However, the cost of early detection can be high and citizen science is a solution to provide long-term and large-scale surveillance for invasive alien species (Johnson et al. 2020). Surveillance of invasive species by citizen scientists is often based on opportunistic reporting of observations by the general public (Pocock et al. 2017), but it would be of benefit to know where it is best to search for recently introduced species in order to focus detection efforts by professionals and citizen scientists. Prior knowledge on the distribution of species from similar countries can be very helpful to predict the likely distribution of invasive species, as has been done for forest pests in the United States (Hudgins et al. 2017).

One of the challenges with citizen science data is that there is an uneven distribution of records, due to uneven coverage of recorders (Isaac and Pocock 2015). This is especially problematic at global scales (Beck et al. 2014; Boakes et al. 2010; Martin et al. 2012), and can depend upon the spatial extent of the projects (Lloyd et al. 2020). However, even at national scales the effect of spatial bias is still apparent with the intensity of records from unstructured citizen science projects often being strongly positively related to human population density, roads and certain habitat types (Boakes et al. 2016; Geldmann et al. 2016; Mair and Ruete 2016; Pernat et al. 2021; Petersen et al. 2021). This bias is affected by aspects of the design and complexity of the citizen science project, and the species being sampled (Boakes et al. 2016; Geldmann et al. 2016; Petersen et al. 2021). There are several methods to account for these biases in analyses, especially where there is data from which to infer absences (Bird et al. 2014; Johnston et al. 2020; Robinson et al. 2018). For example, observations of other species sampled in the same study can be

used as evidence that observers sampled an area but did not find the target species (Phillips et al. 2009), and detecting invasive alien species can be a motivator for people to undertake multispecies sampling (Petersen et al. 2021). However, when unstructured citizen science only targets a single species, such as for an invasive species of concern, then it is harder to account for these biases because it is difficult to distinguish between absence of the species and absence of observers. Potentially, if we can understand these patterns of recording, biases and gaps in recording, then they can inform where to prioritise future sampling (Pocock et al. 2017; Tulloch et al. 2013).

Here we selected the oak lace bug (*Corythucha arcuata* Say, [1832] (Hemiptera: Tingidae) as a model species to explore the influence of citizen science biases for modelling the distribution of an invasive alien species. The oak lace bug is native in North America and was first sampled in Italy (2000) and Turkey (2002) and has spread throughout Europe (Csóka et al. 2020). In the last decade, strong infestations have been reported from Russia (Csóka et al. 2020). The species is able to spread rather quickly as it is often spread passively; records of the oak lace bug have often been made near roads and railways, where all life stages of the insect can be transported by pedestrians, traffic and trains (Csóka et al. 2020; Küçükbaşmacı 2014; Liebhold et al. 2013; Simov et al. 2018), although this could also be due to more observers in these areas. The species can also be passively spread by the wind (Csóka et al. 2020; Küçükbaşmacı 2014). The oak lace bug mainly feeds on oak species (*Quercus* spp.) but is also capable of infesting other types of woody plants, such as elm (*Ulmus* spp.), lime (*Tilia* spp.) and rose (*Rosa* spp.) (Csóka et al. 2020) as spillover from oaks. The damage to the trees is caused by sap-feeding by adults and nymphs, which causes a decrease in photosynthesis (Nikolić et al. 2018), premature defoliation and discolouration of the leaves (Csóka et al. 2020; Paulin et al. 2020) and susceptibility to other pests, diseases and pollution (Küçükbaşmacı 2014). The oak lace bug is regarded as a problem for both forestry and recreation because it can potentially affect the wood production and the aesthetics of the forests (Bălăcenoiu et al. 2021).

The oak lace bug is a species that will potentially gain the attention of the general public because of its strong visual impact on oak trees (Bălăcenoiu et al.

2021). Therefore, citizen scientists can have great value in the early detection of this species, but guidance is required where citizen scientists should make observations for the early detection of the oak lace bug.

The aims of this paper are to test the usefulness of opportunistic citizen science observations for creating a model that predicts where the species is likely to be first detected, and to explore the impact of biases of the citizen science data in the model outputs. We explored the influence of predictor variables on species distribution models that both accounted for the sampling bias of occurrences due to the uneven distribution of citizen science sampling and those that did not. For the model that accounted for this bias, we expected that the predictor variables will be related to the biology and invasion pathways for this species (i.e. Csóka et al. 2020; Küçükbaşmacı 2014; Liebhold et al. 2013), because biases in citizen science are accounted for in the model. With the model that did not explicitly account for this bias, we expected that the variables will be partly confounded with the behaviour of citizen scientists, (Boakes et al. 2016; Petersen et al. 2021), leading to inaccurate estimates of the presence of *C. arcuata*. We further expected that this would show greater uncertainty than the model accounting for recorder bias. We aimed to use the model accounting for recorder bias to make recommendations about designing future early detection with citizen science data for this species. We used the invasion of the oak lace bug in Slovenia as a test case. The first occurrence of the oak lace bug was in 2016 (Jurc and Jurc 2016), which coincided with the start of the citizen science project LIFE ARTEMIS and so the oak lace bug was recorded through that project (Crow et al. 2020). Therefore, the early stage of invasion was potentially well recorded with citizen science, which has not been the case in any other country in which the oak lace bug was introduced. Having the combination of both data of the oak lace bug and other alien species gave us the opportunity to test the influence of biases of recorder behaviour on analyses of invasive species spread.

## Materials and methods

### Area description

Slovenia is a small Central European country (20,273 km<sup>2</sup>) located between Italy, Austria, Hungary and

Croatia. Despite its size, the country is geographically diverse. It is made up of portions of four major European geographic landscapes: the Hungarian plains (Pannonian region) in the east, the karstic Dinaric mountains in the south, and the Alpine region in the north and northwest (Ciglič and Perko 2012); these are divided into six different climatic regions (Kozjek et al. 2017).. Furthermore, Slovenia is an important crossroad for transport in Europe (with major routes from Austria to Croatia and Italy to Hungary), which could be important for the introduction and accidental transport of the oak lace bug (Csóka et al. 2020)]. Oak species, the main host for oak lace bug, are distributed all over the country: *Quercus cerris* is distributed in the west and the east of the country, *Q. petraea* is all over the country, *Q. pubescens* mainly in the west of the country and *Q. robur* mainly in the central part and the east of the country (Jogan et al. 2001).

### Databases

Data on the presence of the oak lace bug were collected over the period 2017–2020 as part of a citizen science activity run through the LIFE ARTEMIS project (LIFE15 GIE/SI/000770) (Crow et al. 2020). In this project, media attention was given to the first observations and possible consequences of the oak lace bug in Slovenia in order to encourage citizen scientists to collect observations of this species with the “Invazivke” information system (Ogris 2020). The “Invazivke” information system is a database for collecting records on non-native species (plants, fungi, insects, and mammals) associated with forests in Slovenia. In total 158 species are in “Invazivke” of which 94 species now have citizen science occurrence data. Records of all species were georeferenced and accompanied by photographic evidence, allowing them to be verified.

For the study we used data for the presence of the oak lace bug and records of other species from observers who also observed oak lace bug from “Invazivke” during 2017–2020 because we knew that they had the ability to identify this species (Arzenšek et al. 2021) and included six additional presence records from a single professional in 2017 when the LIFE ARTEMIS project had only just begun. In order to reduce the impact of clusters of records, we used a 250×250 m grid across Slovenia and randomly

selected one observation per grid cell (hereafter a 'location') per year to give a total of 204 unique location x year combinations. For the constrained pseudo-absence records we selected unique location x year combinations without records of oak lace bug, but with records of any other species recorded in "Invazivke" made by the set of people who had recorded oak lace bug ( $n=1378$ ). We restricted ourselves to this set of people because we were confident that they had the ability to detect oak lace bug if it had been present when they recorded other species. We describe these as pseudo-absences (Phillips et al. 2009) because they were from grid cells without observations of the oak lace bug, but without confirmation that the oak lace bug was absent. For the analysis, we randomly selected a number of constrained pseudo-absence points, i.e. data points of other alien species reported by the same people in the information system "Invazivke", each year that was equal to the number of oak lace bug records that year. We also selected an equal number of random pseudo-absences, i.e. selected from any location without a record of oak lace bug.

A range of predictor variables was used in this study covering physical, human and ecological attributes (Table 1). These variables differed in their spatial resolution and so were unified to the model grid cell dimension of  $250 \times 250$  m ( $n=327,540$ ) covering the whole area of Slovenia. We resampled the elevation from the source 12.5 m spatial resolution to the model cell resolution using the Resample tool of Esri ArcMap 10.6.1 with bilinear interpolation, which calculates the value of each pixel by averaging the values of the surrounding four pixels (GURS 2006). We calculated the slope from a digital elevation model with the Slope tool of ESRI ArcGIS

software and planar method, that calculates slope on a projected flat plane using a 2D Cartesian coordinate system (GURS 2006). We obtained the volume of oak trees ( $m^3$ ) as a proxy for oak tree density, from a forest stand database that is managed by Slovenia Forest Service (ZGS 2020). Almost all oak species in Slovenia were considered, i.e. *Quercus petraea*, *Q. robur*, *Q. pubescens* and *Q. cerris*, because they are the hosts to the oak lace bug. Then, we calculated the shortest distance of each grid cell centroid to forest edges, highways, local roads, and railways using STDistance function of geometry datatype in Microsoft SQL Server 2016 (GURS 2020b). We calculated the area of settlement as an area of buildings in the grid cell using the Land Register of Slovenia (GURS 2020a). Finally, we projected the citizen science data for the oak lace bug and other invasive alien species onto the grid cells, resulting in 204 unique grid cells with oak lace bug records and 1378 with records of other species. Duplicate cells were removed.

#### Statistical analysis

The variables were checked for outliers and multicollinearity using visual inspection (Zuur et al. 2010) and a variance inflation factor using the library "car" (Fox and Weisberg 2019) using the statistical programming language R (R Core Team 2021). The elevation and the surface of urbanized areas, distance to local road, distance to railway, distance to highway, area of the settlement and amount of oak were  $\log_{10} + 1$  transformed in order to reduce the leverage of the extreme values. The variables were visualized with a density plot prepared in the library ggplot2 (Wickham 2016) and described in the results. We then standardized all variables to have a mean of zero

**Table 1** Variables used for the oak lace bug models

| Variable    | Description                                   | Data source  |
|-------------|-----------------------------------------------|--------------|
| x, y        | Coordinates of the centroid of the grid cell  | -            |
| Elevation   | Altitude above sea level (m)                  | GURS (2006)  |
| Slope       | Slope of the terrain (°)                      | GURS (2006)  |
| Oak         | Amount of oaks ( $m^3$ )                      | ZGS (2020)   |
| Forest edge | Shortest distance to a forest edge (m)        | ZGS (2020)   |
| Highway     | Shortest distance to a highway (m)            | GURS (2020b) |
| Local road  | Shortest distance to a local road (m)         | GURS (2020b) |
| Railway     | Shortest distance to a railway (m)            | GURS (2020a) |
| Settlement  | Area of settlement in the grid cell ( $m^2$ ) | GURS (2020a) |

and a standard deviation of one because this gave standardized effect sizes that allowed us to directly compare the importance of the variables (Schielzeth 2010). For the analysis, a generalized linear model (GLM) was used with a binomial error structure.

In total, we ran three species distribution models each of which tested for the effect of these variables against dependent variables of presence and pseudo-absence data. The constrained pseudo-absence model was trained on the presence of oak lace bug and constrained pseudo-absence points: this was to assess the true distribution of the oak lace bug, by accounting for the uneven distribution of citizen science recorders. The random pseudo-absence model was with the presence of oak lace bug and random pseudo-absence points: this assessed the distribution of the oak lace bug, but did not account for the uneven distribution of citizen science recorders. A sizeable difference between random pseudo-absence model and the constrained pseudo-absence model would indicate that the uneven distribution of recorders likely led to biases that affected the results of analysis. A third model was run to describe the spatial biases in the pattern of citizen science reporting. This citizen science reporting model was trained on all citizen science data from the “LIFE ARTEMIS” project in the “Invazivke” database and an equal number of random pseudo-absence points.

We included the independent variables listed in Table 1 in each model. For the constrained pseudo-absence model and the citizen science reporting model, spatial autocorrelation in the residuals was detected with the Moran’s I index. Therefore, the Moran’s Eigenvector was incorporated in these models with the function “ME” from the R package “spatialreg” (Bivand and Piras 2015) to avoid underestimating the magnitude of the parameters.

We built all models using all the predictor variables in Table 1 and evaluated them via cross-validation. Training and validation datasets were selected randomly with a ratio of 80:20. For the training dataset, this was done separately for both random and constrained pseudo-absence (each with 1,103 data points) and presence data (164 data points). For the validation dataset, there were 40 data points for the presence data and 275 pseudo-absence data points (both for the random and constrained pseudo-absence points). In order to test the sensitivity of the models to the input data, we ran the constrained pseudo-absence

model and random pseudo-absence model six times, each with different subsets of pseudoabsences (random or constrained, depending on the model). For the citizen science reporting model there was a larger dataset available for the citizen science data compared to the other models, but to support comparability with the other models, the models were run randomly chosen subsets of the data that were equal in size to the oak lace bug models, and this was replicated six times as with the models described above. To provide a single estimate of the size and significance of each variable from across the six replicates of each of the three models, we used Monte Carlo simulation: for each variable, we simulated 1,000 iterations from the mean and variance from each replicate model, and hence calculated the overall mean and variance across models. We fitted the averaged model on the validation dataset for validation of the models. The area under curve (AUC) in an receiver operating characteristic (ROC) plot was calculated for the validation dataset with the library “PresenceAbsence” (Freeman and Moisen 2008).

We expected that there may be a systematic difference between the constrained and random pseudo-absence models in the model estimates and their precision, and that this might be affected by estimated probability of presence for a location. We tested this by comparing the estimate and its variance from the constrained and random pseudo-absence models for a range of locations. The estimates of probability were not evenly distributed across the range 0–1 so, in order to balance the distribution of points across the estimated values of presence, we randomly selected ten cells from each 0.1 interval of probability of the constrained pseudo-absence model (in total 100 cells: ten randomly selected from the range 0 to 0.1, ten from 0.1 to 0.2, and so on). For each selected cell we obtained the probabilities of presence from the constrained and random pseudo-absence models. We also calculated the precision of these estimates by using Monte Carlo simulation with the model parameters and their variances to generate 10,000 predictions and calculated the inter-quartile range (IQR) of the predictions from each model for each selected cell. IQR was square root arcsine transformed to reduce the bias due to the boundary effect of the proportion data (Gotelli and Ellison 2012) The prediction and transformed IQR was plotted against the prediction from the constrained pseudoabsence model and a “loess”



smoother was applied using the library “ggplot2” (Wickham 2016) to identify differences in variability between the two models.

## Results

Initial observation of the data indicated that the distribution of environmental variables differed between the oak lace bug records and the constrained and random pseudo-absence data (Fig. 2). In particular, the citizen science records (oak lace bug and other species) appeared to be distributed closer to highways, roads, railways and the areas with lower slope than the randomly selected locations. The oak lace bug tended to be recorded further east than other citizen science records.

The constrained pseudo-absence model was expected to be the most accurate description of the variables explaining the oak lace bug distribution because it accounted for spatial biases in the citizen science data. It shows that the oak lace bug was, from the strongest to weakest standardized slopes, significantly more likely to be found where there are more oaks, in the east of the country, closer to highways, at lower elevation and close to railways (Table 2, Fig. 1b, Fig. 2). The year was also an explanatory variable, showing that in 2020 there was a higher probability of recording compared to the reference year 2017, which is probably indicative of the increasing abundance of the species over time, but could also be explained by increased awareness of the species amongst reporters.

The random pseudo-absence model had some broad similarities to the constrained pseudo-absence model in terms of magnitude and significance of many of the variables (Table 2). The six replicates of this model did not show any spatial autocorrelation. However, there were important differences between this and the constrained pseudo-absence model. These were most visible in the map of the predictions (Fig. 2), with higher predicted probability of presence in more accessible locations (near roads and habitation) in the west of Slovenia. These differences indicate the importance of taking account of spatial biases in reporting when analysing these data, although the similarities between the two models indicate that some broad patterns can be identified correctly despite spatial biases in the presence data.

The citizen science reporting model explained the distribution of citizen science records. Although the records will be constrained by the presence of tree pests and diseases and alien plants, we expected that, overall, they provide a reasonable assessment of the distribution of recorders. This model suggests that people are most likely to record in accessible places (from highest to lowest standardized slope): at lower elevations, more in the western part of Slovenia, near transportation, especially local roads, near the forest edge, in places with more human habitation, in places with less sloping ground, close to railways and highways, and in with larger amount of oak trees (Table 2, Fig. 1d, Fig. 2), although in this model only the longitude and distance from local road were statistically significant. This model shows that the difference between the two oak lace bug distribution models (Fig. 1b and c) was explained by the spatial bias in citizen science recording.

The constrained pseudo-absence model’s validation AUC (0.85) was similar to that of the citizen science reporting model (0.83) but considerably lower than that of the random pseudo-absence model (0.95), and this may be because the latter was overfit, suggesting that sampling bias can inflate AUC scores. Finally, we expected that the model estimates and their precision would vary between the constrained pseudo-absence model (i.e. the model that took account of the spatial bias in citizen science reporting) and the random pseudo-absence model. We found that in general, the random pseudo-absence model overestimated the probability compared to the constrained pseudo-absence model (Fig. 3a). Also, uncertainty was lowest for very low and high estimates of probability, but for any level of predicted probability, the random pseudo-absence model had higher uncertainty than the constrained pseudo-absence model (Fig. 3b). This suggests that not taking account of the uneven spatial distribution of citizen science recording leads to predictions that, in this case, are biased to be higher than expected and to have greater uncertainty than expected. Therefore, using information on the distribution of recording effort in modelling is important for accurate predictions of the probability of presence.

**Table 2** Results of the three models predicting the distribution of oak lace bug with constrained and random pseudo-absences, and the distribution of citizen science reports

| Models                                                        | Variables            | Mean    | SE     | P      |     |
|---------------------------------------------------------------|----------------------|---------|--------|--------|-----|
| Constrained pseudo-absence model of oak lace bug distribution | Intercept            | - 2.167 | 0.765  | 0.005  | **  |
|                                                               | log(1 + oak)         | 1.179   | 0.411  | 0.004  | **  |
|                                                               | log(1 + forest edge) | 0.386   | 0.419  | 0.357  |     |
|                                                               | log(1 + highway)     | - 0.778 | 0.365  | 0.033  | *   |
|                                                               | log(1 + localroad)   | - 0.043 | 0.234  | 0.854  |     |
|                                                               | log(1 + railway)     | - 0.488 | 0.248  | 0.049  | *   |
|                                                               | log(1 + settlement)  | 0.271   | 0.275  | 0.325  |     |
|                                                               | elevation            | - 0.739 | 0.334  | 0.027  | *   |
|                                                               | log(1 + slope)       | - 0.399 | 0.304  | 0.189  |     |
|                                                               | longitude            | 0.794   | 0.265  | 0.003  | **  |
|                                                               | Year(2018)           | - 0.944 | 1.036  | 0.362  |     |
|                                                               | Year(2019)           | 1.994   | 0.819  | 0.015  |     |
| Year(2020)                                                    | 3.456                | 0.836   | <0.001 | ***    |     |
| Random pseudo-absence model of oak lace bug distribution      | Intercept            | - 2.898 | 0.858  | 0.001  | **  |
|                                                               | log(1 + oak)         | 1.239   | 0.403  | 0.002  | **  |
|                                                               | log(1 + forest edge) | -0.035  | 0.412  | 0.932  |     |
|                                                               | log(1 + highway)     | -1.037  | 0.262  | <0.001 | *** |
|                                                               | log(1 + localroad)   | -0.583  | 0.360  | 0.106  |     |
|                                                               | log(1 + railway)     | -0.781  | 0.333  | 0.019  | *   |
|                                                               | log(1 + settlement)  | 0.458   | 0.351  | 0.192  |     |
|                                                               | elevation            | -1.231  | 0.821  | 0.134  |     |
|                                                               | log(1 + slope)       | -0.895  | 0.327  | 0.006  | **  |
|                                                               | longitude            | 0.394   | 0.387  | 0.309  |     |
|                                                               | Year(2018)           | 0.156   | 0.867  | 0.857  |     |
|                                                               | Year(2019)           | 2.373   | 0.809  | 0.003  | **  |
| Year(2020)                                                    | 3.369                | 0.992   | 0.001  | **     |     |
| Citizen science reporting model                               | Intercept            | -1.090  | 0.495  | 0.028  | *   |
|                                                               | log(1 + oak)         | 0.129   | 0.258  | 0.617  |     |
|                                                               | log(1 + forest edge) | -0.411  | 0.275  | 0.135  |     |
|                                                               | log(1 + highway)     | -0.340  | 0.292  | 0.245  |     |
|                                                               | log(1 + localroad)   | -0.498  | 0.234  | 0.033  | *   |
|                                                               | log(1 + railway)     | -0.332  | 0.266  | 0.213  |     |
|                                                               | log(1 + settlement)  | 0.376   | 0.309  | 0.224  |     |
|                                                               | elevation            | -0.747  | 0.426  | 0.080  |     |
|                                                               | log(1 + slope)       | -0.340  | 0.254  | 0.180  |     |
|                                                               | longitude            | -0.689  | 0.233  | 0.003  | **  |
|                                                               | Year(2018)           | 1.412   | 0.681  | 0.038  | *   |
|                                                               | Year(2019)           | 1.476   | 0.540  | 0.006  | **  |
| Year(2020)                                                    | 0.836                | 0.832   | 0.315  |        |     |

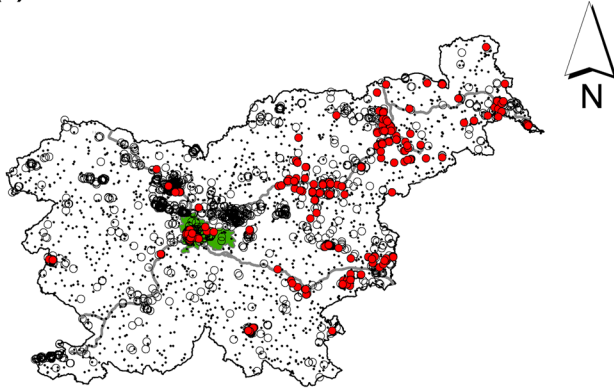
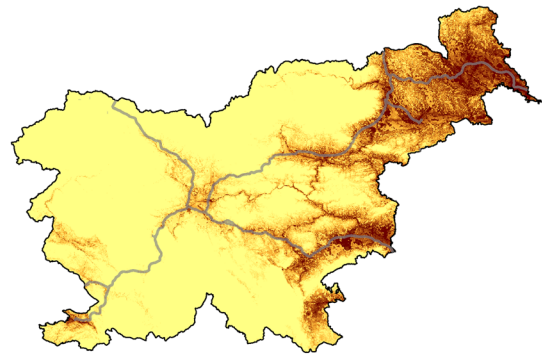
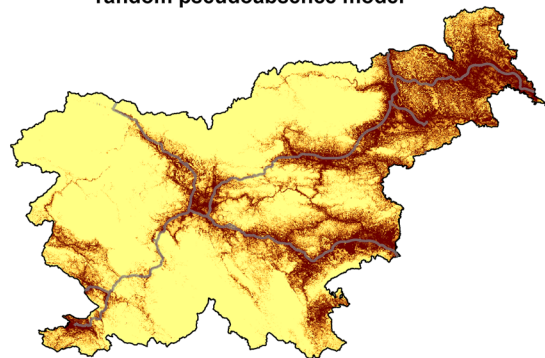
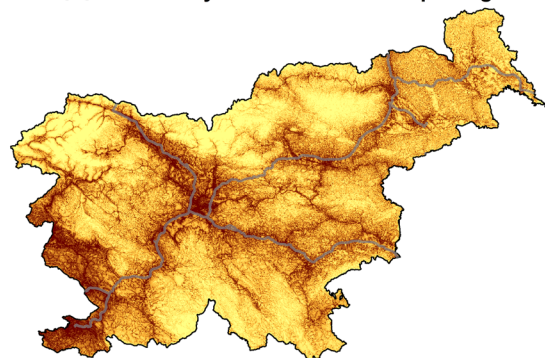
The mean and SE were obtained (via Monte Carlo sampling) from six replicates of each model with different selections of pseudo-absences/absences as described in the text. \*  $P < 0.05$ , \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$ . All variables were standardized with mean and standard deviation

## Discussion

The distribution of the oak lace bug in Slovenia

Here we showed that a national, multi-year citizen science dataset was effective for assessing the

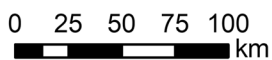
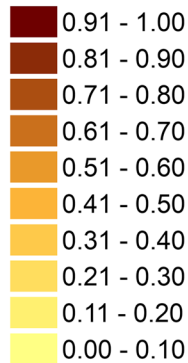
increasing distribution of an invasive tree pest, the oak lace bug, in Slovenia. In common with many citizen science projects, this one was unstructured (i.e. 'opportunistic'), meaning that people could contribute when and where they chose. Here we modelled habitat suitability for the oak lace bug with citizen

**(a) Location of sites****(b) Presence of oak lace bug - constrained pseudoabsence model****(c) Presence of oak lace bug - random pseudoabsence model****(d) Probability of citizen science reporting**

## Legend

- Oak lace bug
- Citizen science
- Randomly selected absence points
- Highway
- ✚ Ljubljana
- 🇸🇮 Slovenia

## Probability



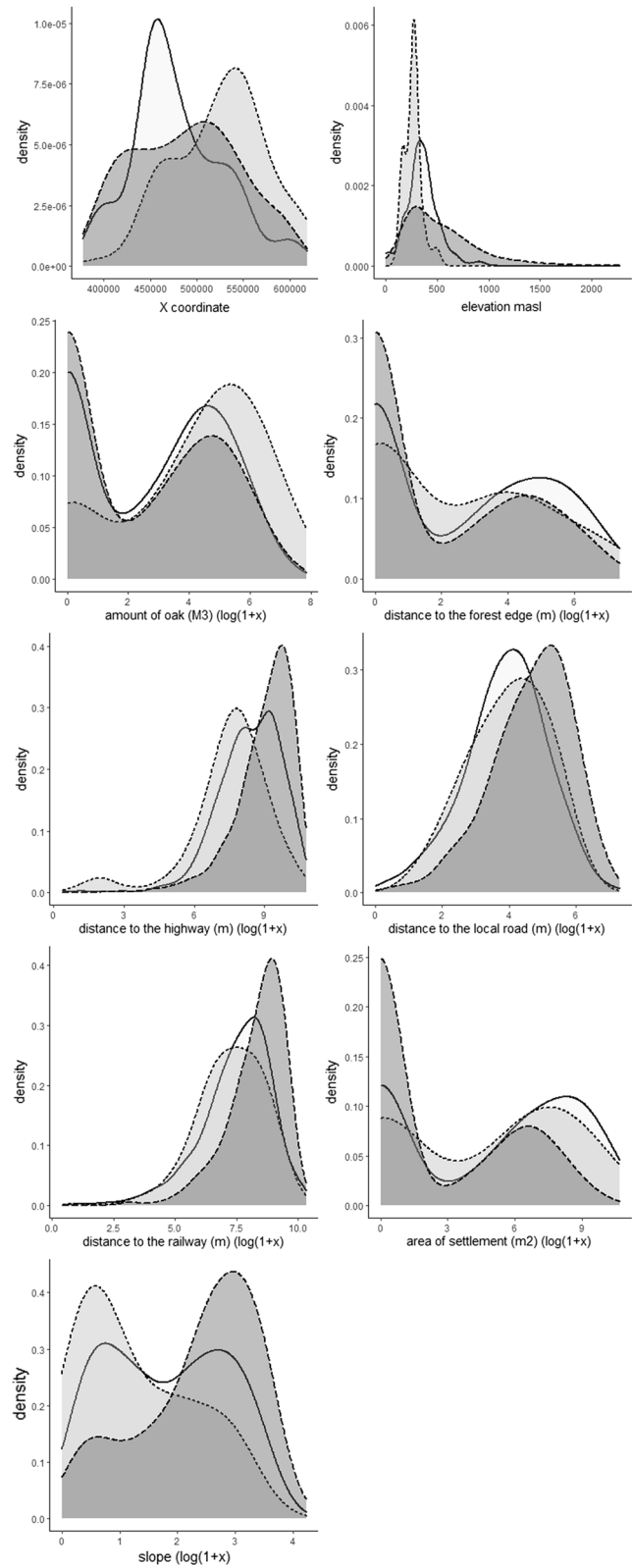
Datasources: OLB and CS data ([www.invazivke.si](http://www.invazivke.si)), RFMU (SFS)

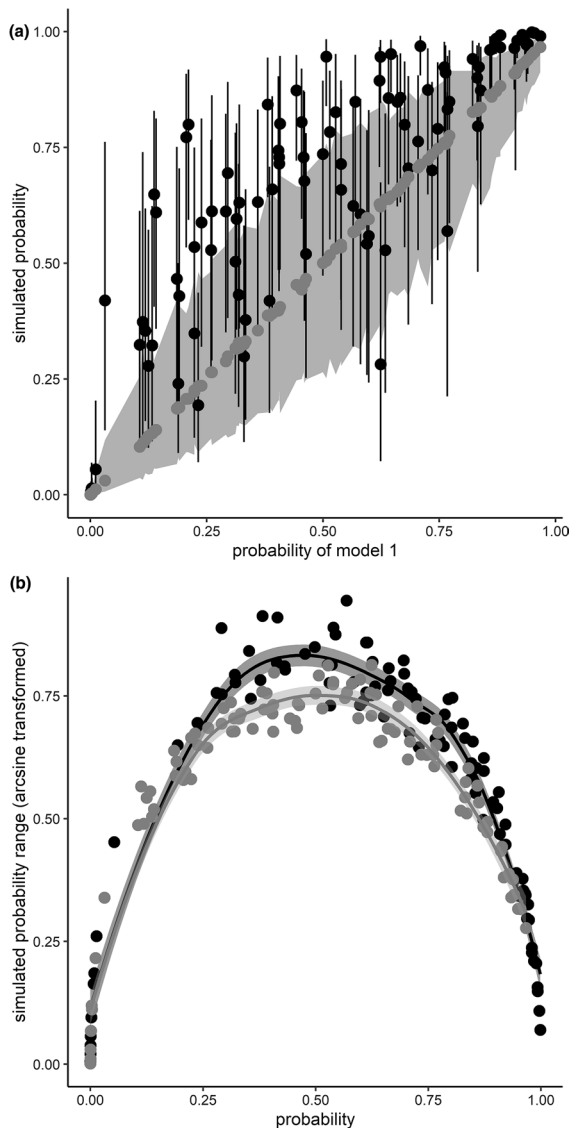
**Fig. 1** The observed and predicted distribution of oak lace bug and citizen science recording effort. **(a)** Maps with the presence of oak lace bug (OLB: filled circles), data from other citizen science projects on tree pests and diseases (CS: open circles) and random points used for modelling (small dots). Predicted distribution of the oak lace bug in 2020 based on a

species distribution model that used **(b)** constrained pseudo-absences, and **(c)** random pseudo-absences. **(d)** The predicted distribution of citizen science reporting model. Legend: Darker colours shows higher probability of the presence. RFMU = regional forest management unit



**Fig. 2** Density plots of environmental variables, showing the differences in the distribution of the presence of oak lace bug (light grey, short dashed line), other citizen science data (white, continuous line) and the random points (dark grey, long dashed line)





**Fig. 3** Simulated probability of the presence of oak lace bug and its uncertainty from 1000 randomly selected cells. Figure 3a shows the correlation between the simulated probability of constrained pseudo-absence model and the random pseudo-absence model, and the probability of constrained pseudo-absence model. The grey dots and ribbon are the simulated probability mean and the variability between the first and third quartile of the constrained pseudo-absence model. The black dots and lines are the simulated probability mean and the variability between the first and third quartile of the random pseudo-absence model. Figure 3b shows the simulated variability between the first and third quartile association with the probability of the constrained pseudo-absence model (black line) and the random pseudo-absence model (grey line)

science data on occurrences, taking the uneven distribution of recorders into account, to predict where oak lace bug occurred. The observed locations of the oak lace bug generally matched the predictions from our species distribution model (constrained pseudo-absence model in Table 2; Fig. 1a and b). However, it appeared to be under-recorded in the far east of Slovenia, relatively over-reported around the large city of Ljubljana, and there was a region in the south west of Slovenia where the oak lace bug was predicted to be present but with no records yet, indicating likely suitability of the region for invasion by this species.

The model showed the relationship of oak lace bug with predictor variables, specifically that the species was more likely to be present near highways and railways, in places with high oak density and at lower altitudes (constrained absence model in Table 2). Although not much is known about its natural dispersal, adults of the closely related plane lace bug (*C. ciliata*) do not fly long distances (Wu and Lui 2016), so the association with major travel routes indicates that it spreads to other locations via traffic and railroads (Csóka et al. 2020; Küçükbaşmacı 2014; Liebhold et al. 2013; Simov et al. 2018), probably via accidental transport of oak logs. Indeed the first observation in Slovenia was along a railroad (Jurc and Jurc 2017), and one early observation was at a rest stop along the highway near Maribor (pers. comm. M. de Groot).

The oak lace bug first arrived in the east of Slovenia (Jurc and Jurc 2017), from the Balkans and Hungary, where large outbreaks had already been found (Csóka et al. 2020; Jurc and Jurc 2017), and spread westwards, so latitude was an important predictor of current distribution. However, several citizen science observations have been already made in the west of the country. Populations do not seem to be well established there yet, but it seems likely that populations will establish there in the future, with our model indicating regions most susceptible to early establishment (Fig. 2b).

We also found that the amount of oak positively affects the presence of the oak lace bug. This is not surprising because the species mainly, although not exclusively, uses oaks as its host (Csóka et al. 2020). Outbreaks of oak lace bug could be mitigated by increased tree species diversity as was shown also for other herbivores (Jactel et al. 2017), although it remains a question whether changing woodland structure would decrease the likelihood of outbreaks.

In Slovenia, oak is more prevalent at lower altitudes (San-Miguel-Ayanz et al. 2016) but even taking oak abundance into account, altitude was also an important predictor; it remains uncertain whether oak lace bug is restricted by temperature, as is the case for many species (Ratte 1984), or whether it will spread to these less accessible places over time.

We note that this model shows only where the species will be found in this stage of its invasion; a thorough assessment of habitat suitability for the species can be obtained when it invades the whole area and is closer to dynamic equilibrium.

#### Best use of citizen science in species distribution modelling

In this project, we had the benefit of using presence records from other species in the same citizen science project as pseudo-absences in our dataset. This was important because the citizen science recording effort was spatially biased (Model 3 in Table 2). The pattern of citizen science recording was strongly influenced by accessibility of sites to people, as shown by the importance (absolute size of standardised regression coefficient) of variables such as proximity to settlements, local roads and railways, proximity to forest edges, and elevation (Table 2). Several other studies have found that citizen science recording is strongly influenced by factors related to accessibility at both broad scales, e.g. proximity to towns, and local scales, e.g. proximity to roads (Boakes et al. 2010; Isaac and Pocock 2015; Mair and Ruete 2016; Petersen et al. 2021).

By comparing the results of the models with constrained pseudoabsences to that with random pseudoabsences (i.e. the models that did and did not take account of recorder bias, respectively), we showed that not taking recorder effort into account led to biased model outputs. The random pseudoabsence model (Fig. 2c) had higher predicted presence near settlements in the centre and west of Slovenia, compared to the constrained pseudoabsence model (Fig. 2b). Consideration of differences in the model variables (Table 2) also showed that the model with random pseudoabsences found that being near a forest edge had higher probability of presence. All these differences mirror the distribution of citizen science recording effort (Fig. 2d). In general, the model with random pseudo-absences tended to have

higher predicted presence for any specific location (Fig. 3a) and, for any value of predicted presence, tended to be less precise (greater uncertainty; Fig. 3b), which was because it did not account for the spatial bias in the distribution of citizen science recorders.

All these results show the importance of taking the distribution of citizen science recording effort into account when analysing citizen science data of invasive species. However, gaining a true understanding of the distribution of potential recording effort is very difficult. In the past researchers have ignored the problem or used proxies such as other people recording similar species (Pocock et al. 2017). Here, we took pseudo-absence points from the records of other tree pests and diseases from citizen science, so using it as a proxy for recorder effort. However, this was not a perfect measure of ‘absence’: the other species were from many different species, many were not insects and some occurred at different seasons to the oak lace bug. Also, it is possible that there may be some places (e.g. pristine forests) where people looked but did not record any of the species, so recording effort would be under-estimated there. One solution to this problem would be to ask people to record absences. This would work best with a more structured sampling approach, but the motivation to record an absence is very different to the motivation to record a presence, and structured sampling is inefficient, and arguable unethical (Pocock et al. 2020), for early detection where the probability of detection is low. In addition, paucity of any data from less accessible locations (Fig. 2d) could reduce the statistical power to identify some biologically-informative variables in the model with constrained pseudoabsences.

Many citizen science projects only record presence data, and there is no information on recording effort, so it is not possible to run a model with constrained pseudoabsences. Although there are important differences between our two models of oak lace bug distribution (Fig. 2 and 3, Table 2), the general patterns of distribution are similar. This indicates that presence-only data can be used for species distribution modelling, although great care should be taken when interpreting the results of analysis. Ultimately, including the distribution of potential recording in unstructured citizen science is a vital advance that will support the good analysis of citizen science data for the early detection of invasive species.

## Implications of our findings for early detection of invasive species

As well as considering what the results tell us about the distribution of oak lace bug, they also help us consider how to improve citizen science for the early detection of this species. The finding that oak lace bug was most likely to be found along highways, even once biases in recorder effort were taken into account, highlights that these types of sites should be targeted for early detection of this species. Indeed, citizen science records were also more likely near highways (Citizen science reporting model in Table 2), there is a valuable convergence between the places where people choose to record and the places where they need to look for early detection of this species.

Planning for early detection should include the use of citizen science (Pocock et al. 2017), given the growing importance of citizen science in this field (Roy et al. 2018). Therefore, when designing citizen science for oak lace bug in areas where it is not yet present, such as western Slovenia or neighbouring countries, it would be valuable to recommend recorders focus on recording from stands of oaks near highways, in an effort to target early detection of the species (Morissette et al. 2020). Conversely, where the species is becoming established, it could be more valuable to gather records from places where we have least information, e.g. further away from settlements and at higher elevations.

Overall, citizen science is a valuable approach to invasive species early detection and monitoring. This method did not exclude biases and challenges in the use of the data but, as we have shown, with care these can be overcome to make the most of the data and so honour the valuable contribution from members of the public. We recommend that understanding the distribution of the potential recording effort from citizen science is vital to benefit effectively from this opportunity. It helps us properly use the results for construction of unbiased distribution models of the invasive species and so understand predictors of the distribution. This will help us to better target recording effort, whether through citizen science, e.g. via local publicity campaigns and targeted information on where to look, or through professional monitoring to fill gaps in the citizen science recording effort.

**Acknowledgements** We would like to thank the volunteers who submitted data to the Invazivke information system.

**Authors' contributions** Conceptualization: Maarten de Groot, Michael Pocock; data management: Nikica Ogris; Formal analysis and investigation: Maarten de Groot; visualization: Nikica Ogris, Maarten de Groot; Writing—original draft preparation: Maarten de Groot, Mark van der Meij, Nikica Ogris, Michael Pocock; Writing—review and editing: Maarten de Groot, Mark van der Meij, Nikica Ogris, Michael Pocock.

**Funding** The authors want to acknowledge the LIFE ARTEMIS project (LIFE15 GIE/SI/000770), which was co-financed by the LIFE programme, Ministry of Environment and Spatial planning of the Republic of Slovenia, Slovenian Research Agency and the municipality of Ljubljana, and the research core group “Forest Biology, Ecology and Technology” (P4-0107) and the COST action Alien CSI “Increasing understanding of alien species through citizen science” (COST Action CA17122). MJOP was supported by the Natural Environment Research Council award number NE/R016429/1 as part of the UK-SCAPE programme delivering National Capability.

**Availability of data and material** The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

**Code availability** Not applicable.

## Declarations

**Conflict of interest** None of the authors had a conflict of interests.

**Ethics approval** Not applicable.

**Consent to participate** All authors agree with the participation.

**Consent for publication** All authors agree with the publication.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Arzenšek B, de Groot M, Šneberger B, Zidar S, Brglez A, Zagorac N, Malovrh J, Bogovič M, Jurman E, Ogris N, Groznik E, Kostanjevec M, Strniša A, Potisk Z, Smolnikar P, Laganiš J, Janevič D, Piškur B, Držaj A, Gradišnik Mirt M, Trajber D, Anonymous, Cigan A, Žveplan A, Rozman J, Šram A, Štrbac D, Orožim M, Simčič A, Sarjaš A, Gorjup E, Maltar Ž, Foltin S, Stiplošek A, Zavrtnik Z (2021) Records of invasive alien species in Slovenia from 2017–2020. Ogris N., de Groot, M, Kutnar, L., Marinšek, A. (eds). Invazivke - Central electronic information system for invasive alien species in Slovenia, [www.invazivke.si](http://www.invazivke.si). Gozdarski inštitut Slovenije. LIFE ARTEMIS (LIFE15 GIE/SI/000770). 15.10.2021
- Bălăcenoiu F, Japelj A, Bernardinelli I, Castagneyrol B, Csóka G, Glavendekić M, Hoch G, Hrašovec B, Krajter Ostoic S, Paulin M, Williams D, Witters J, de Groot M (2021) *Corythucha arcuata* (Say, 1832) (Hemiptera, Tingidae) in its invasive range in Europe: perception, knowledge and willingness to act in foresters and citizens. *NeoBiota* 69:133–153. <https://doi.org/10.3897/neobiota.69.71851>
- Battisti A, Benvegnù I, Colombari F, Haack RA (2014) Invasion by the chestnut gall wasp in Italy causes significant yield loss in *Castanea sativa* nut production. *Agric for Entomol* 16:75–79. <https://doi.org/10.1111/afe.12036>
- Beck J, Böller M, Erhardt A, Schwanghart W (2014) Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Eco Inform* 19:10–15. <https://doi.org/10.1016/j.ecoinf.2013.11.002>
- Bird TJ, Bates AE, Lefcheck JS, Hill NA, Thomson RJ, Edgar GJ, Stuart-Smith RD, Wotherspoon S, Krkosek M, Stuart-Smith JF, Pecl GT, Barrett N, Frusher S (2014) Statistical solutions for error and bias in global citizen science datasets. *Biol Cons* 173:144–154. <https://doi.org/10.1016/j.biocon.2013.07.037>
- Bivand R, Piras G (2015) Comparing Implementations of Estimation Methods for Spatial Econometrics. *J Stat Softw* 63:1–36
- Boakes EH, McGowan PJK, Fuller RA, Chang-qing D, Clark NE, O'Connor K, Mace GM (2010) Distorted Views of Biodiversity: Spatial and Temporal Bias in Species Occurrence Data. *PLoS Biol* 8:e1000385. <https://doi.org/10.1371/journal.pbio.1000385>
- Boakes EH, Gliozzo G, Seymour V, Harvey M, Smith C, Roy DB, Haklay M (2016) Patterns of contribution to citizen science biodiversity projects increase understanding of volunteers' recording behaviour. *Sci Rep* 6:33051. <https://doi.org/10.1038/srep33051>
- Ciglič R, Perko D (2012) Slovenia in geographical typifications and regionalizations of Europe. *Geografski Vestnik* 84:23–37
- Clavero M, García-Berthou E (2005) Invasive species are a leading cause of animal extinctions. *Trends Ecol Evol* 20:110. <https://doi.org/10.1016/j.tree.2005.01.003>
- Crow P, Perez-Sierra A, Kavčič A, Lewthwaite K, Kolšek M, Ogris N, Piškur B, Kus Veenvliet J, Zidar S, Sancisi-Frey S, de Groot M (2020) Using Citizen Science to monitor the spread of tree pests and diseases: Outcomes of two projects in Slovenia and the UK. *Management of Biological Invasions* 11:703–719
- Csóka G, Hirka A, Mutun S, Glavendekić M, Mikó Á, Szócs L, Paulin M, Eötvös CB, Gáspár C, Csepelényi M, Szénási Á, Franjevič M, Gninenko Y, Dautbašić M, Muzejinović O, Zúbrík M, Netoiu C, Buzatu A, Bălăcenoiu F, Jurc M, Jurc D, Bernardinelli I, Streito J-C, Avtziš D, Hrašovec B (2020) Spread and potential host range of the invasive oak lace bug [*Corythucha arcuata* (Say, 1832) – Heteroptera: Tingidae] in Eurasia. *Agric for Entomol* 22:61–74. <https://doi.org/10.1111/afe.12362>
- Fox J, Weisberg S (2019) An {R} Companion to Applied Regression. Sage, Thousand Oaks {CA}. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>
- Freeman EA, Moisen G (2008) PresenceAbsence: An R Package for Presence Absence Analysis. *J Stat Softw* 23:1–31
- Geldmann J, Heilmann-Clausen J, Holm TE, Levinsky I, Markussen B, Olsen K, Rahbek C, Tøttrup AP (2016) What determines spatial bias in citizen science? Exploring four recording schemes with different proficiency requirements. *Divers Distrib* 22:1139–1149. <https://doi.org/10.1111/ddi.12477>
- Gotelli NJ, Ellison A (2012) A primer of ecological statistics. Sinauer Associates
- GURS (2006) Digital elevation model 12,5 m. Surveying and Mapping Authority of the Republic of Slovenia, Slovenia TSaMAotRo
- GURS (2020a) Building cadastre. Surveying and Mapping Authority of the Republic of Slovenia, Ljubljana
- GURS (2020b) Consolidated cadastre of public infrastructure. Surveying and Mapping Authority of the Republic of Slovenia, Ljubljana
- Hudgins EJ, Liebhold AM, Leung B (2017) Predicting the spread of all invasive forest pests in the United States. *Ecol Lett* 20:426–435. <https://doi.org/10.1111/ele.12741>
- Hulme PE (2009) Handbook of alien species in Europe Springer. The Netherlands, Dordrecht
- Isaac NJB, Pocock MJO (2015) Bias and information in biological records. *Biol J Lin Soc* 115:522–531. <https://doi.org/10.1111/bij.12532>
- Jactel H, Bausch J, Boberg J, Bonal D, Castagneyrol B, Gardiner B, Gonzalez-Olabarria JR, Koricheva J, Meurisse N, Brockerhoff EG (2017) Tree Diversity Drives Forest Stand Resistance to Natural Disturbances. *Current Forestry Reports* 3:223–243. <https://doi.org/10.1007/s40725-017-0064-1>
- Jogan N, Bacic T, Frajman B, Leskovar I, Naglic D, Podobnik A, Rozman B, Strgulc-Krajsek S, Trcak B (2001) Materials for the Atlas of Flora of Slovenia. Center for Carthography of Fauna and Flora, Ljubljana
- Johnson BA, Mader AD, Dasgupta R, Kumar P (2020) Citizen science and invasive alien species: An analysis of citizen science initiatives using information and communications technology (ICT) to collect invasive alien species observations. *Global Ecology and Conservation* 21:e00812. <https://doi.org/10.1016/j.gecco.2019.e00812>
- Johnston A, Moran N, Musgrove A, Fink D, Baillie SR (2020) Estimating species distributions from spatially biased citizen science data. *Ecol Model* 422:108927. <https://doi.org/10.1016/j.ecolmodel.2019.108927>



- Jurc D, Jurc M (2017) The first record and the beginning the spread of oak lace bug, *Corythucha arcuata* (Say, 1832) (Heteroptera: Tingidae), in Slovenia. *Šumarski List* 141:488–488. <https://doi.org/10.31298/sl.141.9-10.5>
- Kozjek K, Dolinar M, Skok G (2017) Objective climate classification of Slovenia. *Int J Climatol* 37:848–860. <https://doi.org/10.1002/joc.5042>
- Küçükbasmacı İ (2014) Two new invasive species recorded in Kastamonu (Turkey): Oak lace bug [*Corythucha arcuata* (Say, 1832)] and sycamore lace bug [*Corythucha ciliata* (Say, 1832)] (Heteroptera: Tingidae). *Journal of Entomology and Nematology* 6:104–111. <https://doi.org/10.5897/JEN2014.0102>
- Liebholt AM, McCullough DG, Blackburn LM, Frankel SJ, Von Holle B, Aukema JE (2013) A highly aggregated geographical distribution of forest pest invasions in the USA. *Divers Distrib* 19:1208–1216. <https://doi.org/10.1111/ddi.12112>
- Lloyd TJ, Fuller RA, Oliver JL, Tulloch AI, Barnes M, Steven R (2020) Estimating the spatial coverage of citizen science for monitoring threatened species. *Global Ecology and Conservation* 23:e01048. <https://doi.org/10.1016/j.gecco.2020.e01048>
- Lowe S, Browne M, Boudjelas S, De Poorter M (2000) 100 of the world's worst invasive alien species: a selection from the global invasive species database vol 12. Invasive Species Specialist Group, Auckland, New Zealand
- Mair L, Ruete A (2016) Explaining Spatial Variation in the Recording Effort of Citizen Science Data across Multiple Taxa. *PLoS ONE* 11:e0147796. <https://doi.org/10.1371/journal.pone.0147796>
- Marcolin E, Pividori M, Colombari F, Manetti MC, Pelleri F, Conedera M, Gehring E (2021) Impact of the Asian gall wasp *Dryocosmus kuriphilus* on the radial growth of the European chestnut *Castanea sativa*. *J Appl Ecol* 58:1212–1224. <https://doi.org/10.1111/1365-2664.13861>
- Martin LJ, Blossley B, Ellis E (2012) Mapping where ecologists work: biases in the global distribution of terrestrial ecological observations. *Front Ecol Environ* 10:195–201. <https://doi.org/10.1890/110154>
- Morisette JT, Reaser JK, Cook GL, Irvine KM, Roy HE (2020) Right place. Right time. Right tool: guidance for using target analysis to increase the likelihood of invasive species detection. *Biol Invasions* 22:67–74. <https://doi.org/10.1007/s10530-019-02145-z>
- Nikolić N, Pilipovic A, Drekić M, Kojić D, Poljaković-Pajnik L, Orlović S, Arsenov D (2018) Physiological responses of Pedunculate oak (*Quercus robur* L.) to *Corythucha arcuata* (Say, 1832) attack. *Archives of Biological Sciences*:58–58 doi:<https://doi.org/10.2298/ABS180927058N>
- Ogris N (2020) Web portal “Invazivke”. Slovenian Forestry Institute, Ljubljana
- Paulin M, Hirka A, Eötvös C, Csaba G, Fürjes-Mikó Á, Csóka G (2020) Known and predicted impacts of the invasive oak lace bug (*Corythucha arcuata*) in European oak ecosystems – a review. *Folia Oecologica* 47:131–139. <https://doi.org/10.2478/foecol-2020-0015>
- Pernat N, Kampen H, Ruland F, Jeschke JM, Werner D (2021) Drivers of spatio-temporal variation in mosquito submissions to the citizen science project ‘Mückenatlas.’ *Sci Rep* 11:1356. <https://doi.org/10.1038/s41598-020-80365-3>
- Petersen TK, Speed JDM, Grøtan V, Austrheim G (2021) Species data for understanding biodiversity dynamics: The what, where and when of species occurrence data collection. *Ecological Solutions and Evidence* 2:e12048. <https://doi.org/10.1002/2688-8319.12048>
- Phillips SJ, Dudík M, Elith J, Graham CH, 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. <https://doi.org/10.1890/07-2153.1>
- Pocock MJO, Roy HE, Fox R, Ellis WN, Botham M (2017) Citizen science and invasive alien species: Predicting the detection of the oak processionary moth *Thaumetopoea processionea* by moth recorders. *Biol Cons* 208:146–154. <https://doi.org/10.1016/j.biocon.2016.04.010>
- Pocock MJO, Marzano M, Bullas-Appleton E, Dyke A, de Groot M, Shuttleworth CM, White R (2020) Ethical dilemmas when using citizen science for early detection of invasive tree pests and diseases. *Management of Biological Invasions* 11:720–732
- R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
- Ratte HT (1984) Temperature and Insect Development. In: Hoffmann KH (ed) *Environmental Physiology and Biochemistry of Insects*. Springer, Berlin, Heidelberg,
- Robinson OJ, Ruiz-Gutierrez V, Fink D (2018) Correcting for bias in distribution modelling for rare species using citizen science data. *Divers Distrib* 24:460–472. <https://doi.org/10.1111/ddi.12698>
- Roy HE, Groom Q, Adriaens T, Agnello G, Antic M, Archambeau A-S, Bacher S, Bonn A, Brown P, Brundu G, López BC, Cleary M, Cogălniceanu D, de Groot M, De Sousa T, Deidun A, Essl F, Fišer Pečnikar Ž, Gazda A, Gervasini E, Glavendekic MM, Gigot G, Jelaska SD, Jeschke JM, Kaminski D, Karachle PK, Komives T, Lapin K, Lucy F, Marchante E, Marisavljevic D, Marja R, Martín Torrijos L, Martinou A, Matosevic D, Mifsud CM, Motiejūnaitė J, Ojaveer H, Pasalic N, Pekárik L, Per E, Pergl J, Pesic V, Pocock M, Reino L, Ries C, Rozyłowicz L, Schade S, Sigurdsson S, Steinitz O, Stern N, Teofilovski A, Thorsson J, Tomov R, Tricarico E, Trichkova T, Tsiamis K, van Valkenburg J, Vella N, Verbrugge L, Véték G, Villaverde C, Witzell J, Zenetos A, Cardoso AC (2018) Increasing understanding of alien species through citizen science (Alien-CSI). *Research Ideas and Outcomes* 4:e31412
- San-Miguel-Ayanz J, de Rigo D, Caudullo G, Durrant T, Mauri A, Tinner W, Ballian D, Beck P, Birks H, Eaton E, Enescu C, Pasta S, Popescu I, Ravazzi C, Welk E, Abad Viñas R, Azevedo J, Barbati A, Barredo J, Zecchin B (2016) *European Atlas of Forest Tree Species*. Publications Office of the European Union, Luxembourg
- Schielzeth H (2010) Simple means to improve the interpretability of regression coefficients. *Methods Ecol Evol* 1:103–113. <https://doi.org/10.1111/j.2041-210X.2010.00012.x>
- Seebens H, Blackburn TM, Dyer EE, Genovesi P, Hulme PE, Jeschke JM, Pagad S, Pyšek P, Winter M, Arianoutsou



- M, Bacher S, Blasius B, Brundu G, Capinha C, Celesti-Grapow L, Dawson W, Dullinger S, Fuentes N, Jäger H, Kartesz J, Kenis M, Kreft H, Kühn I, Lenzner B, Liebhold A, Mosena A, Moser D, Nishino M, Pearman D, Pergl J, Rabitsch W, Rojas-Sandoval J, Roques A, Rorke S, Rossinelli S, Roy HE, Scalera R, Schindler S, Štajerová K, Tokarska-Guzik B, van Kleunen M, Walker K, Weigelt P, Yamanaka T, Essl F (2017) No saturation in the accumulation of alien species worldwide. *Nat Commun* 8:14435. <https://doi.org/10.1038/ncomms14435>
- Seidl R, Klöner G, Rammer W, Essl F, Moreno A, Neumann M, Dullinger S (2018) Invasive alien pests threaten the carbon stored in Europe's forests. *Nat Commun* 9:1626. <https://doi.org/10.1038/s41467-018-04096-w>
- Simov N, Grozeva S, Langourov M, Georgieva M, Mirchev P, Georgiev G (2018) Rapid expansion of the Oak lace bug *Corythucha arcuata* (Say, 1832) (Hemiptera: Tingidae) in Bulgaria. *Historia Naturalis Bulgarica* 27:51–55
- Spatz DR, Zilliacus KM, Holmes ND, Butchart SHM, Genovesi P, Ceballos G, Tershy BR, Croll DA (2017) Globally threatened vertebrates on islands with invasive species. *Sci Adv* 3:e1603080. <https://doi.org/10.1126/sciadv.1603080>
- Tulloch AIT, Mustin K, Possingham HP, Szabo JK, Wilson KA (2013) To boldly go where no volunteer has gone before: predicting volunteer activity to prioritize surveys at the landscape scale. *Divers Distrib* 19:465–480. <https://doi.org/10.1111/j.1472-4642.2012.00947.x>
- Vila M, Espinar JL, Hejda M, Hulme PE, Jarošik V, Maron JL, Pergl J, Schaffner U, Sun Y, Pyšek P (2011) Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. *Ecol Lett* 14:702–708
- Wickham H (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York
- ZGS (2020) Forest stands map. Slovenia Forest Service. Ljubljana.
- Zuur AF, Ieno EN, Elphick CS (2010) A protocol for data exploration to avoid common statistical problems. *Methods Ecol Evol* 1:3–14. <https://doi.org/10.1111/j.2041-210X.2009.00001.x>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.