Mapping species distributions in two weeks using citizen science

Seirian Sumner[1*], Peggy Bevan[1], Adam G. Hart[2] & Nick J.B. Isaac[1,3]

1. **Centre for Biodiversity and Environment Research, University College London, Gower Street, London, WC1E 6BT**
2. **School of Natural and Social Science, University of Gloucestershire, Cheltenham, GL50 4AZ**
3. **Centre for Ecology and Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, OX10 8BB**

***Corresponding author: s.sumner@ucl.ac.uk**

Keywords: Insect, wasps, species distributions, bias, citizen science

Abstract

1. Ecological citizen science initiatives are growing in popularity with the increasing realisation of the potential for occurrence records to contribute information on biodiversity. However, citizen science data are justifiably criticised for misidentification, uneven sampling, incomplete detection or selective reporting.
2. Here, we test the accuracy of citizen science data for UK social wasp (Vespinae) species' distributions. We compared data collected over two weeks by members of the public setting out baited traps across the UK and sending captured specimens for expert identification (1294 locations; 6680 wasps; 3 dominant species *Vespula vulgaris* (44%), *V. germanica* (44%), and *Vespa crabro* (6%)), with a four-decade long-term dataset established by the Bees, Wasps and Ants Recording Society (BWARS).
3. The citizen science data were significantly less spatially biased than the long-term data, but they were more urban-biased. Species distribution modelling showed that, for *Vespa crabro*, just two weeks of citizen science generated coverage comparable to more than four decades of expert recording.
4. Overall, we show that citizen science can be an extremely powerful and robust method for mapping insect diversity and distributions. We suggest that cautious combination of citizen science data with long-term expert surveying could be a highly reliable method for monitoring biodiversity.

Introduction

Systematic biodiversity surveys are expensive and time-consuming. Consequently, there has been a growing realisation of the potential for occurrence records to contribute information on biodiversity (Tingley and Beissinger, 2009). Recording schemes by expert amateur recorders (Isaac and Pocock, 2015) and, more recently, non-expert "citizen scientists” can generate useful data (Pocock *et al.*, 2015). Mass participation citizen science projects appealing to the public can be especially useful because they can take advantage of wide geographical coverage and potentially thousands of observers. By recruiting help from the public, projects have been able to study a range of taxa and ecological phenomena (Pocock *et al.*, 2017). However, the validity of such projects has been questioned; for example, most citizen scientists lack the expertise to identify species (Lewandowski and Specht, 2015; Maldonado *et al.*, 2015a), records tend to be opportunistic with uneven spatial coverage (Silvertown, 2009) and sometimes the lack of formal protocols makes it difficult to understand why some species were not recorded (Pocock *et al.*, 2017). These problems result in data that are patchy (Maldonado *et al.*, 2015b) and often biased towards charismatic groups attractive to the public, e.g. Big Garden Bird Watch (Lewandowski and Specht, 2015). Despite taxonomic biases, there are examples of citizen science approaches on non-charismatic taxa such as spiders (Campbell and Engelbrecht, 2018; Hart, Nesbit and Goodenough, 2018; Wang *et al.*, 2018) and beetles (Zapponi *et al.*, 2017), and producing valid ecological insights (Pocock *et al.*, 2017; Hart *et al.*, 2018). Overall, it is not always straightforward to assess the validity of citizen science data since high-quality datasets for direct comparison are rarely available (Kamp *et al.*, 2016; Dennis *et al.*, 2017). Of course, the lack of high-quality datasets is one of the principle motivating factors for professional scientists to work with citizens (Silvertown *et al.*, 2009; Strien, Swaay and Termaat, 2013).

To investigate the validity of citizen science data for species distribution mapping, the UK-based Big Wasp Survey (BWS: www.bigwaspsurvey.org) focussed on an uncharismatic but well-known and relatively easily distinguished taxon of social wasps (genera *Vespula, Dolichovespula* and *Vespa*) with a simple protocol that removed misidentification problems; participants captured wasps over one-week using standardised baited traps and sent them for expert identification. We compare the citizen science data generated over two weeks from these collections with data for the same species generated by experienced recorders of the Bees, Wasps and Ants Recording Society (BWARS) over a 46-year period. This approach allowed us to investigate the extent to which a time-focused citizen science project can generate useful data on species' distributions and to compare the usefulness of citizen science data for constructing species distribution models with data gathered by experienced recorders.

Methods

2.1. BWS Data

The Big Wasp Survey (BWS) ran from 26th August to the 10th September 2017. Advertised through national media, participants left out a wasp trap filled with beer or orange juice (Fig. 1; instructions at www.bigwaspsurvey.org) for 7 days. Trap contents with metadata (dates, bottle volume and colour, liquid used) were sent for identification by trained volunteers (blind-checked by PB and SS). It was possible for the study to use trained volunteers for identification because of the small number of potential target species (10 in total, including the newly arrived invasive species *Vespa velutina*). The centroid of the post codes for each trap location were converted to OSGB grid references.

2.2. Recording Scheme Data

We compared BWS data to BWARS records spanning 1970-2016 for the three most common species recorded by BWS (*Vespula vulgaris, Vespula germanica, Vespa crabro*). The BWARS records are generated in an *ad hoc* manner. BWARS is a subscription-based society (http://www.bwars.com/) whose membership largely consists of expert amateur naturalists. The BWARS data consist of opportunistic observations collected in an *ad hoc* manner. BWARS coordinates a network of regional experts who have responsibility for validating and verifying all records that are submitted to the society. The verification process includes consultation between recorder and expert in cases where the identification is questionable. The BWARS data consist of 7226 unique records of social wasps from 3378 1km grid cells (henceforth 'monads') in 838 10x10km grid cells (henceforth 'hectads').

2.3 Spatial bias

To test the extent to which BWS and BWARS data were spatially biased we compared the number of 1km cells in England, Wales and Scotland with the number that would be expected if sample locations were distributed randomly (Northern Ireland and Isle of Man excluded due to small numbers) using Chi-Square tests. We then conducted the same comparison on the 10 aggregate landcover classes from the CEH Landcover map 2015 (Rowland *et al.*, 2017) by classifying each km^2 according to its dominant landcover type.

Finally, we measured the spatial aggregation of each dataset by comparing cumulative frequency distribution of nearest neighbour distances (*r*) for each dataset with the null expectation. We compared this with the distribution of *r* expected if samples were unbiased by generating 99 sets of random points with sample size equal to the focal data set. To quantify deviations from random we summed differences between observed and expected nearest neighbour distances at each value of *r* to give a summary statistic, *res_diff*.

For these spatial bias tests, we conducted the analysis for BWS using both the full set of monads and the subset in which at least one wasp was recorded. The former (“BWS traps”) reflects the distribution of effort, whilst the latter (“BWS wasps”) reflects the distribution of wasp records, so is more directly comparable with the BWARS data.

2.4. Species Distribution Models (SDM)

We explored the information content of each dataset by fitting species distribution models (SDMs) to the records for the hornet, *Vespa crabro*. This species is well-suited for an investigation of this nature because it has a clear range margin in central Britain, unlike the two *Vespula* species, which are both distributed throughout Great Britain. We constructed SDMs based on BWS (n=100 monads), BWARS (n=1895) and a third dataset comprising the union (n=1995). For each we sampled pseudoabsences from sites where either *V. vulgaris* or *V. germanica* were recorded but not *V. crabro* (BWS n=424; BWARS n=8488, union=8939). The number of pseudoabsences was set equal to the number of presence records. We fitted the distribution of *V. crabro* for both datasets using four covariates: the fitted distribution using a GLM and a GAM using the zoon package (Golding *et al.*, 2017) in R with five predictor variables: mean temperature, summer precipitation (Fick, 2017), human population density (Reis *et al.*, 2017), the areas of arable land and area of semi-natural habitat (Rowland *et al.*, 2017).

Results

3.1 BWS data

Of the 2377 people registered, 1294 people (54.4%) submitted their results from 1279 unique postcodes (1251 monads; 776 hectads). Of these, 551 traps (42.5%) contained wasps, representing 548 unique postcodes (543 monads; 441 hectads). Traps were concentrated around major conurbations (Fig. 1 top panel) and sparser sampling in rural areas, extreme regions (e.g. northern Scotland, west Wales) and national parks (e.g. New Forest, Lake District).

A total of 6680 wasps from 551 traps were identified (180 wasps from 39 traps were too badly decomposed to identify). The three dominant species were *V. vulgaris* (44%; n=2942 wasps; n=407 traps), *V. germanica* (44%; n=2974 wasps; n=251 traps) and *V. crabro* *(*representing 6%; n=395 of the total identified catch; n=100 traps; Fig. 2). These species are known to have long colony cycles in the UK, with workers abundant in late summer. We detected a few individuals of the species with shorter colony cycles, which have typically come to their seasonal end at the time of sampling (*Vespula rufa:* n=2 in 2 traps; *Dolichovespula media:* n=9 in 6 traps). We found that traps were more likely to capture wasps when baited with orange juice than with beer ($\chi 2=7.0$, P=0.007); counter-intuitively, we found that smaller bottles were more likely to capture wasps than larger bottles ($\chi 2=48.3$, P<0.001). We found no effect of bottle colour ($\chi 2=4.23$, P=0.12) on wasp capture.

3.2. Spatial bias

All three datasets were spatially biased, but BWS traps and BWS wasps were both more evenly distributed across the UK than BWARS data (Fig 1, Table 1). All datasets had a higher than expected number of records from England, but this effect is stronger in BWARS data than BWS (Chi-square test (χ^2 nation); $p<0.001$; Table 1). The BWS data - in a single survey window of just 7 days across a two-week period - provided data for over half the number of hectads obtained from 46 years of BWARS data (Fig 2), thus illustrating the power of citizen science for coarse mapping of species distributions.

Both BWS and BWARS were biased with respect to land cover type, but the effect differed among datasets. Both BWS and BWARS are biased towards Built-up areas and grassland, but the effect is far stronger in BWS; in addition, BWARS is strongly biased towards monads dominated by Broadleaf woodland (Table 2).

3.3 Species distribution models

Species distributions generated by BWS and BWARS datasets were similar (Fig 3). The difference between the two different model types (GAM vs GLM), measured by AUC, was virtually identical to the difference between the models created using the BWS and BWARS data (Table 3), suggesting that the two are relatively similar. A combined output of the GLM and GAM models (union) appears identical to those of the BWARS models, suggesting the model is relatively unchanged by the addition of the BWS occurrences, although BWS outperforms both the BWARS and union datasets in most other metrics (Kappa, omission rate and Specificity score). The BWARS model suggested a consistently high occurrence of *Vespa crabro* across southern and central England, whereas BWS (but not BWARS) showed a high occurrence probability in only the most southern parts of the country (Fig 3).

Discussion

We provide strong evidence that citizen scientists can generate remarkably high-quality data within very short timescales. Indeed, in this study a single week of nationwide citizen science sampling generated data with spatial coverage equivalent to more than four decades of expert recording from the UK's current 'gold standard' wasp recording scheme. In this, the BWS is not alone; a Japanese citizen science study of bee distribution also performed better than models created with records from the Global Biodiversity Information Faculty (Suzuki-Ohno *et al.*, 2017).

Critical to this success, BWS data were collected using a standardised method that took advantage of the ubiquity and enthusiasm of the public even for uncharismatic species like wasps, resulting in large sample sizes and high geographical coverage. Moreover, identification to species level, always a potential weak link in citizen science projects, were undertaken by experts. Not only did this result in less geographic bias than established recording schemes (although still with some bias towards urban areas), but also species distribution models' performance of citizen science data produced rapid results.

Despite the success of the citizen science approach, there were issues. A bias towards gardens, and towards centres of high population density, resulted in an overall urban bias and distinct omissions from regions of low population density which would be difficult to rectify even with focussed recruitment. On the other hand, expert surveyors could be deployed to such under-reported regions. Data from the two sampling methods may therefore be complementary, but still omit remote rural environments, farmland and national parks.

Although the emphasis of our study was to examine the robustness of citizen science data, our analyses also revealed some new insights into Vespine biology and distributions in the UK. Firstly, the BWS model also suggested that urban areas may be less suitable for *V. crabro* than more rural areas, and this is despite the higher sampling effort in urban areas. Our results also reveal interesting differences in relative abundance between the two common species. *V vulgaris* was recorded in nearly twice as many trap locations as *V. germanica* (407 vs 251), yet the total number of records for the two species was virtually identical (2942 vs 2974). This suggests that *V. germanica* may be more locally abundant than *V. vulgaris*, or have larger mean colony size, at least during 2017. Bonanza "wasp years" and wasp-poor years have been recorded for both species with potentially two or seven year cycles of yearly abundance (Archer, 1985). The explanation for annual fluctuations in wasp abundances remain elusive, both weather and nest density have been implicated (Lester *et al.*, 2017); long-term annual sampling through BWS may help understand this better, and thus help manage these insects for their ecosystem services and to minimise annoyance for the public (Sumner *et al*., 2018).

The BWS and BWARS datasets differ not only in the protocol and sampling bias, but also in the length of time over which the records were collected. The span of the BWARS data includes a period during which *Vespa crabro* is known to have expanded its distribution in central-southern England since the 1970s (BWARS, 2018). These factors make formal comparison of the SDMs difficult, but nonetheless our study shows that rapidly-collected citizen science data have real value.

In our analysis, we simply compared SDMs from each dataset and from the merged datasets. New “integrated” SDMs account for different observation processes underpinning each dataset, although current implementations (Dorazio, 2014) generally assume that one dataset lacks any spatial bias. However, these developments, and the demonstrable value of snapshots from citizen science, suggest a future in which species are monitored by a suite of methods, each with complementary strengths and weaknesses. A specific strength of BWS is the opportunity to re-engage the same citizen scientists year on year, generating replicated sampling at the same locations. Such data would facilitate the study of how land-use and local climatic fluctuations affect wasp populations, diversity and distributions.

Figure Legends

Figure 1. Cumulative frequency distributions of nearest neighbour distances for wasp records from the Big Wasp Survey (top panel) and the Bees, Wasps and Ants Recording Society (BWARS, bottom panel). The y-axis is the cumulative frequency; the x-axis is the nearest neighbour distance, in metres (note different scale of the two panels). Black lines show the observed distributions, red lines are the expected distributions (if all points were distributed randomly) and the grey area delimits the range of expected values from 99 simulations. Insets show the spatial distribution of sample points (BWS: n = 543 1 km^2 grid cells; BWARS: n = 3378).

Figure 2: Records of the three dominant species caught in the Big Wasp Survey 2017: European hornet (*Vespa crabro* 101 traps), common yellow-jacket wasp (*Vespula vulgaris*; 407 traps) and the German wasp (*Vespula germanica*; 251 traps).

Figure 3: Species distribution maps for the European hornet, *Vespa crabro.* Each map is a combination of a different data set (the 'Big Wasp Survey' (BWS) or the 'Bees, Wasps and Ants Recording Society (BWARS), or a union of the two) and a different model type (general linear model (GLM) or general additive model (GAM). The scale with each map is a measure of habitat suitability based on land cover type and human population density. The output of the unionised models (combination of BWS & BWARS) is very similar to the BWARS models. For ΔAUC values (model performance), see Table 3.

Table 1: Spatial bias in recorder effort of BWS and BWARS. The column headed 'BWS traps' refers to the full dataset; the column headed 'BWS wasps' refers to the subset where at least one of the three common species was captured. Numbers in the first five rows refer to traps where wasps were recorded for BWS, and square kilometres where wasps were recorded for BWARS. Both datasets have higher than expected number of samples in England, but to a stronger degree in BWARS (* = $p<0.001$). Conversely, BWS is more highly biased by land cover (χ^2 land cover). *res_diff* is a measure of aggregation, where 0=random distribution: BWS samples are more evenly distributed across the whole of the UK than BWARS.

Table 2. The level of sampling in different landcover types of BWS and BWARS data compared to the actual proportions of land cover types in the UK. The ratio column refers to the ratio of observed:expected number of monads dominated by each landcover class.

Table 3. Performance metrics for species distribution models for the hornet, *Vespa crabro*.

References

Archer, M.E. (1985) Population dynamics of the social wasps Vespula vulgaris and Vespula germanica in England, *Journal of Animal Ecology*, **54**, 473–485. doi: 10.2307/4492.

Bees, Wasps and Ants Recording Society. <http://www.bwars.com/wasp/vespidae/vespinae/vespa-crabro> 13th August 2018.

Campbell, H. & Engelbrecht, I.A.N. (2018) The Baboon Spider Atlas – using citizen science and the “ fear factor ” to map baboon spider (Araneae: Theraphosidae) diversity and distributions in Southern Africa. *Insect Conservation and Diversity*, **11**, 143–151. doi: 10.1111/icad.12278.

Dennis, E.B., Morgan, B.J.T., Brereton, T.M., Roy, D.B. & Fox, R. (2017) Using citizen science butterfly counts to predict species population trends, *Conservation Biology*, **31,** 1350-1361. doi: 10.1111/cobi.12956.

Dorazio, R.M. (2014) Accounting for imperfect detection and survey bias in statistical analysis of presence-only data, *Global Ecology and Biogeography*, **23,** 1472–1484. doi: 10.1111/geb.12216.

Fick, S.E. (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas, *International Journal of Climatology,* **37**, 4302–4315. doi: 10.1002/joc.5086.

Golding, N., August, T.A., Lucas, T.C.D., Gavaghan, D.J., van Loon, E.E. & McInerny, G. (2017) The zoon r package for reproducible and shareable species distribution modelling, *Methods in Ecology and Evolution*, **9**, 260-268. doi: 10.1111/2041-210X.

Hart, A.G., Hesselberg, T., Nesbit, R. & Goodenough, A.E. (2018) The spatial distribution and environmental triggers of ant mating flights: using citizen-science data to reveal national patterns, *Ecography*, **41**, 877–888. doi: 10.1111/ecog.03140.

Hart, A.G., Nesbit, R. & Goodenough, A.E. (2018) Spatiotemporal Variation in House Spider Phenology at a National Scale Using Citizen Science, *Arachnology*, **17**, 331–335.

Isaac, N.J.B. & Pocock, M.J.O. (2015) Bias and information in biological records, *[in special issue: Fifty years of the Biological Records Centre] Biological Journal of the Linnean Society*, **115**, 522–531. doi: 10.1111/bij.12532.

Kamp, J., Oppel, S., Heldbjerg, H., Nyegaard, T. & Donald, P.F. (2016) Unstructured citizen science data fail to detect long-term population declines of common birds in Denmark, *Diversity and Distributions,* **22**, 1024–1035. doi: 10.1111/ddi.12463.

Lester, P.J., Haywood, J., Archer, M.E. & Shortall, C.R. (2017) The long-term population dynamics of common wasps in their native and invaded range, *Journal of Animal Ecology*, 86, 337-347. doi: 10.1111/1365-2656.12622.

Lewandowski, E. & Specht, H. (2015) Influence of volunteer and project characteristics on data quality of biological surveys, **29**, 713–723. doi: 10.1111/cobi.12481.

Maldonado, C., Molina, C.I., Zizka, A., Persson, C., Taylor, C.M., Albán, J., Chilquillo, E., Rønsted, N. & Antonelli, A. (2015) Estimating species diversity and distribution in the era of Big Data: to what extent can we trust public databases?, *Global Ecology and Biogeography*, **24**, 973-984. doi: 10.1111/geb.12326.

Pocock, M.J.O., Roy, H.E., Preston, C.D. & Roy, D.B. (2015) The Biological Records Centre: A pioneer of citizen science, *Biological Journal of the Linnean Society*, **115**, 475–493. doi: 10.1111/bij.12548.

Pocock, M.J.O., Roy, H.E., Fox, R., Ellis, W.N. & Botham, M. (2017) Citizen science and invasive alien species: Predicting the detection of the oak processionary moth Thaumetopoea processionea by moth recorders, *Biological Conservation*, **208**, 146–154. doi: 10.1016/j.biocon.2016.04.010.

Pocock, M.J.O., Tweddle, J.C., Savage, J., Robinson, L.D. & Roy, H.E. (2017) The diversity and evolution of ecological and environmental citizen science, *PLOS ONE*, 12, e0172579. doi: 10.1371/journal.pone.0172579.

Reis, S., Liska, T., Steinle, S., Carnell, E., Leaver, D., Roberts, E., Vieno, M., Beck, R. & Dragosits, U. (2017) UK Gridded Population 2011 based on Census 2011 and Land Cover Map 2015. *NERC Environmental Information Data Centre*. doi: 0995e94d-6d42-40c1-8ed4-5090d82471e1

Rowland, C.S., Morton, R.D., Carrasco, L., McShane, G., O'Neil, A.W. & Wood, C.M. (2017) Land Cover Map 2015 (1km percentage aggregate class, GB), *NERC Environmental Information Data Centre*. doi: 10.5285/7115bc48-3ab0-475d-84ae-fd3126c20984.

Silvertown, J. (2009) A new dawn for citizen science, *Trends in Ecology and Evolution*, **24**, 467–471. doi: 10.1016/j.tree.2009.03.017.

Strien, A.J. van, Swaay, C.A.M. van & Termaat, T. (2013) Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models, *Journal of Applied Ecology*, **50**, 1450–1458. doi: 10.1111/1365-2664.12158.

Sumner, S., Law, G. & Cini, A. (2018) Why we love bees and hate wasps, *Ecological Entomology*, **43**, 836-845. doi: https://doi.org/10.1111/een.12676.

Suzuki-Ohno, Y, Yokoyama, J., Nakashizuka, T. & Kawata, M. (2017) Utilization of photographs taken by citizens for estimating bumblebee distributions, *Scientific Reports*, **7**, 1–11. doi: 10.1038/s41598-017-10581-x.

Tingley, M.W. & Beissinger, S.R. (2009) Detecting range shifts from historical species occurrences: new perspectives on old data, *Trends in Ecology & Evolution*, **24**, 625–33. doi: 10.1016/j.tree.2009.05.009.

Wang, Y., Casajus, N., Buddle, C., Berteaux, D. & Larrivee, M. (2018) Predicting the distribution of poorly documented species, Northern black widow (*Latrodectus variolus*) and Black purse-web spider (*Sphodros niger*), using museum specimens and citizen science data, *PLOS ONE,* **13**, e0201094.

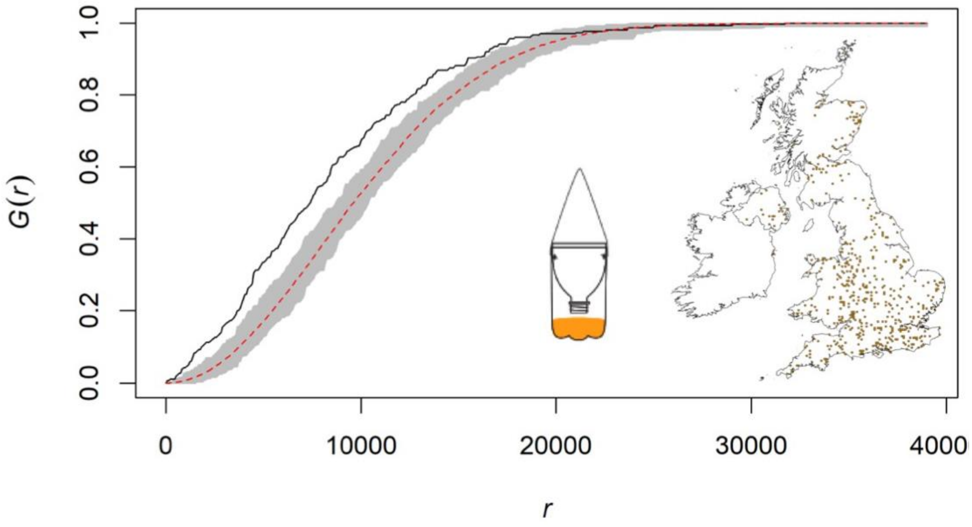
Zapponi, L. *et al.* (2017) Citizen science data as an efficient tool for mapping protected saproxylic beetles, *Biological Conservation*, **208**, 139-145. doi: 10.1016/j.biocon.2016.04.035.

Acknowledgements

We thank the committee and members of BWARS (the UK Bees, Wasps and Ants Recording Society) for access to their data holdings. We are particularly grateful to Stuart Roberts for support and advice. A big thank you to the many members of the public that took part in the BWS! We thank London's Natural History Museum, London Natural History Society members (especially John Lock) and UCL students for sorting, especially Susannah

Warnock. We also thank BWS Webmaster Cris Thompson for all his assistance, and the Royal Entomological Society for funding the postage of samples collected by the public. SS and AH conceived the idea and ran the BWS; SS & PB identified samples; NI & PB conducted analyses. All authors wrote the paper. Data (Appendix 1) are available at www.BigWaspSurvey.org/data. The Royal Entomological Society funded postage for sample return. There are no competing interests to declare. This study was approved by the UCL Ethics Committee.

Figure 1.


1.0
0.8
0.6
0.4
0.2
0.0
G(r)
0
10000
20000
30000
4000
r

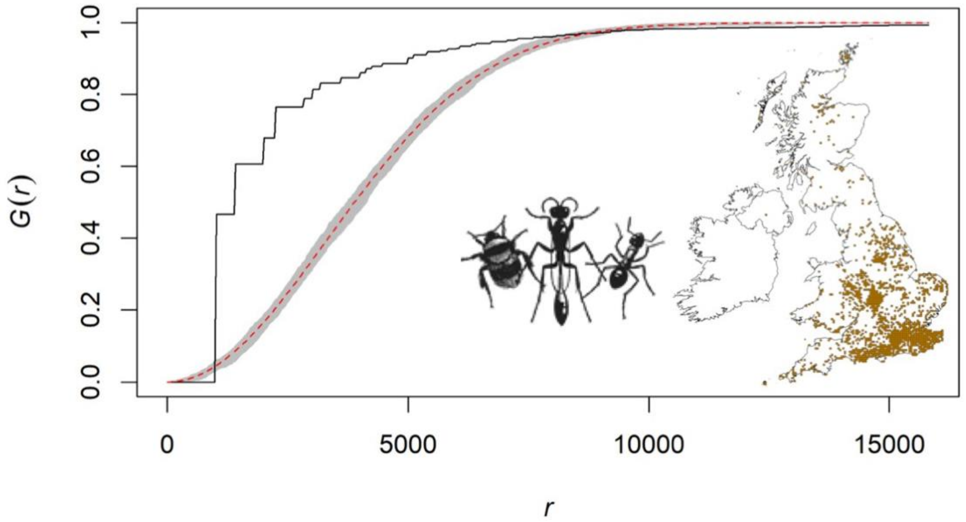

1.0
0.8
0.6
0.4
0.2
0.0
G(r)
0
5000
10000
15000
r

Figure 2

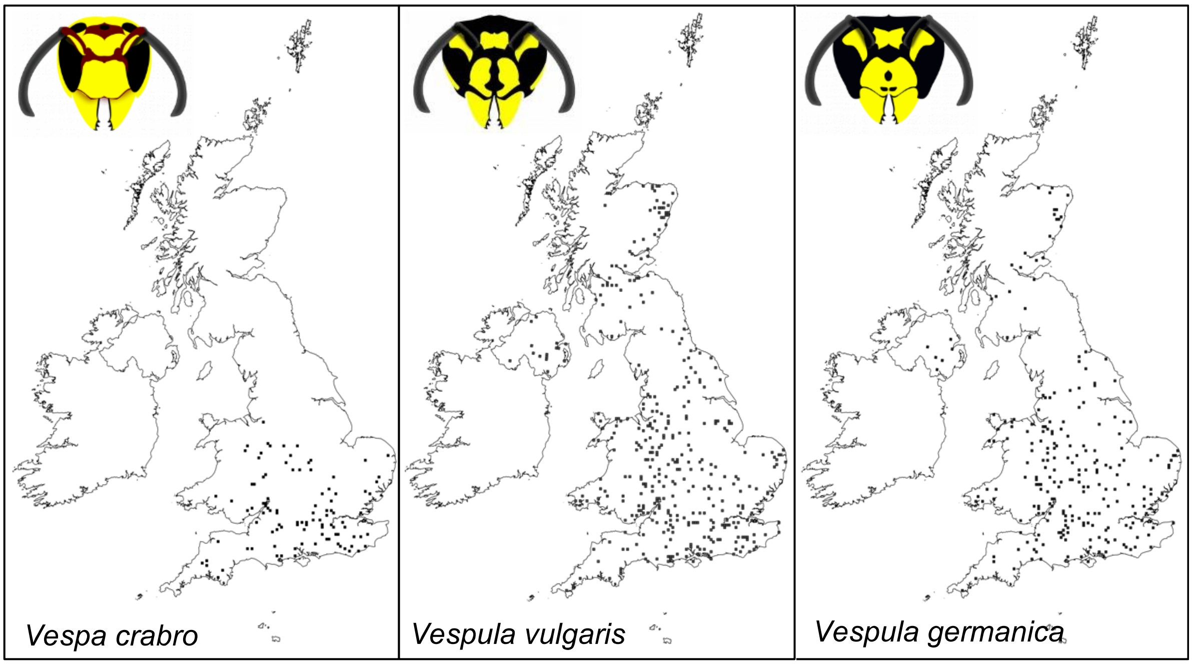
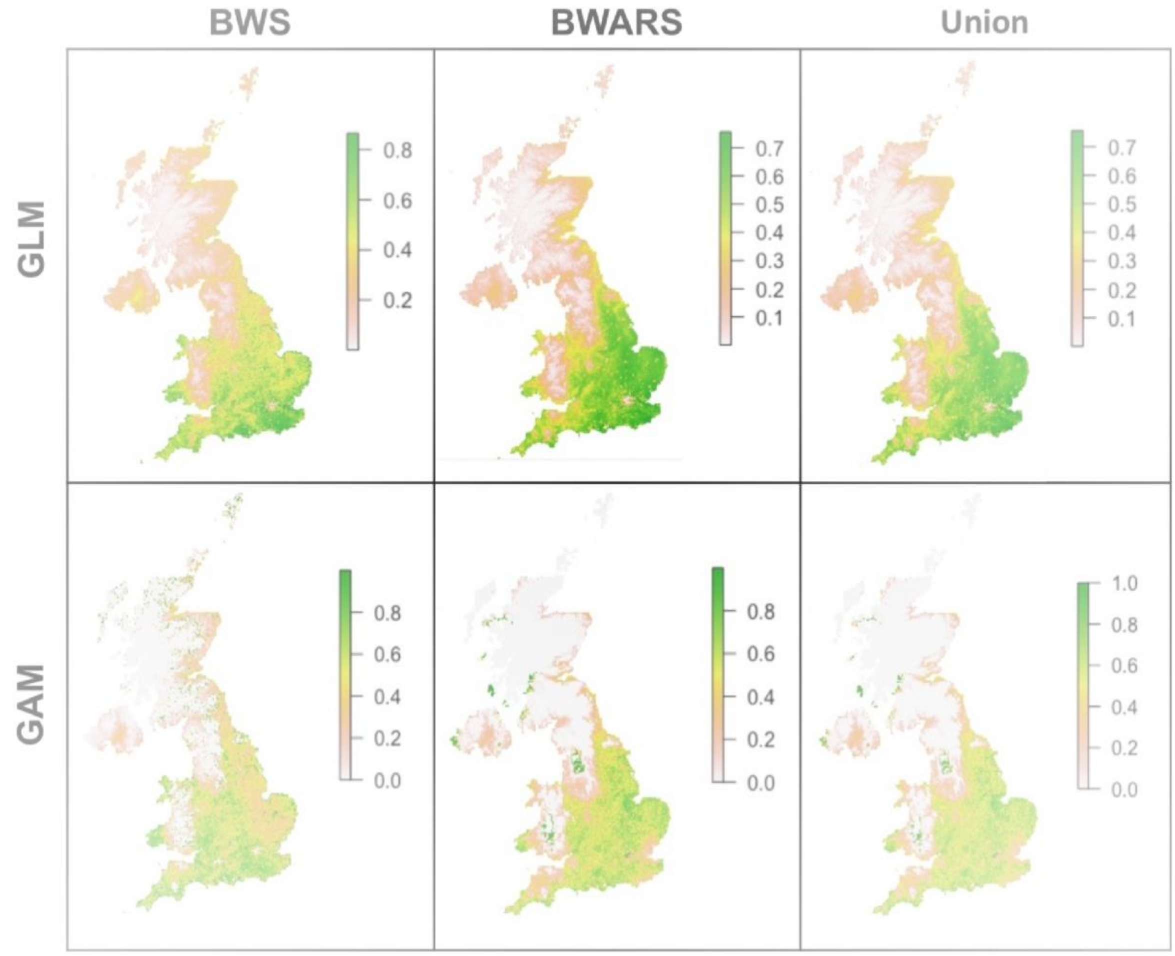


Figure 3:



BWS
BWARS
Union
GLM
GAM
0.8
0.6
0.4
0.2
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.8
0.6
0.4
0.2
0.0
0.8
0.6
0.4
0.2
0.0
1.0
0.8
0.6
0.4
0.2
0.0

Table 1

	BWS traps	BWS wasps	BWARS
Total monads	1279	548	5399
England	1034	392	5137
Wales	85	58	98
Scotland	133	78	163
other	27	20	1
X^2 Nation	395*	100*	3,472*
X^2 Landcover	3,296*	566*	2,748*
res_diff	38.08	26.71	64.84

Table 2

	BWARS		BWS - Traps		BWS - Wasps	
Landcover Class	monads	ratio	monads	ratio	monads	ratio
Broadleaf woodland	455	5.47	25	0.83	16	1.24
Coniferous woodland	147	0.73	7	0.10	4	0.13
Arable	756	0.87	205	0.65	119	0.88
Improved grassland	1087	1.04	337	0.89	202	1.25
Semi-natural grassland	81	0.26	11	0.10	4	0.08
Mountain, heath, bog	76	0.15	3	0.02	0	0
Saltwater	16	0.65	4	0.44	1	0.26
Freshwater	25	1.15	3	0.38	2	0.59
Coastal	67	0.93	9	0.34	3	0.27
Built-up areas and grassland	659	2.98	620	7.72	171	4.99

Table 3

Dataset	Model	AUC	Kappa	Omissions	Sensitivity	Specificity
BWS	GLM	0.737	0.370	0.260	0.740	0.630
	GAM	0.804	0.460	0.230	0.770	0.690
BWARS	GLM	0.666	0.237	0.259	0.741	0.496
	GAM	0.734	0.345	0.209	0.791	0.553
Union	GLM	0.667	0.243	0.261	0.739	0.504
	GAM	0.732	0.340	0.216	0.784	0.555