Article:

Hicks, H., Lambert, J., Pywell, R. et al. (5 more authors) (2021) Characterizing the environmental drivers of the abundance and distribution of Alopecurus myosuroides on a national scale. Pest Management Science, 77 (6). pp. 2726-2736. ISSN 1526-498X

https://doi.org/10.1002/ps.6301

Research Article

Received: 7 April 2020 Revised: 30 December 2020 Accepted article published: 26 January 2021 Published online in Wiley Online Library:

(wileyonlinelibrary.com) DOI 10.1002/ps.6301

Characterizing the environmental drivers of the abundance and distribution of *Alopecurus myosuroides* on a national scale

Helen Hicks,[a] James Lambert,[b,c] Richard Pywell,[b,c] Lucy Hulmes,[b,c] Sarah Hulmes,[b,c] Kevin Walker,[d] Dylan Z Childs[b,c] and Robert P Freckleton[b,c*]

Abstract

BACKGROUND: Arable weeds threaten farming and food production, impacting on productivity. Large-scale data on weed populations are typically lacking, and changes are frequently undocumented until they reach problem levels. Managing the future spread of weeds requires that we understand the factors that influence current densities and distributions. In doing so, one of the challenges is to measure populations at a large enough scale to be able to accurately measure changes in densities and distributions. Here we analyse the density and distribution of a major weed (*Alopecurus myosuroides*) on a large scale. Our objectives were to (i) develop a methodology for rapid measurement of occurrence and abundance, (ii) test hypotheses about the roles of soils and climate variation in determining densities, and (iii) use this information to identify areas in which occurrence could increase in the future.

RESULTS: Populations were mapped through England over 4 years in 4631 locations. We also analysed UK atlas data published over the past 50 years. Densities of populations show significant interannual variability, but historical data show that the species has spread. We find significant impacts of soil and rainfall on densities, which increase with the proportion of heavy soils, but decrease with increasing rainfall. Compared with independent atlas data we found that our statistical models provide good predictions of large-scale occupancy and we provide maps of current and potential densities.

CONCLUSION: Models of spread highlight the localised nature of colonisation, and this emphasises the need for management to limit dispersal. Comparisons of current, historical and potential distributions suggest sizeable habitable areas in which increases in abundance are still possible.

Supporting information may be found in the online version of this article.

Keywords: weed; population model; species distribution; generalized additive mode

1 INTRODUCTION

The large-scale distributions of plants result from factors operating at a hierarchy of scales. Within populations, rates of birth and death, and how these covary with the environment and density drive local abundances.[1–3] Dispersal permits colonisation of new habitats[4] and ultimately determines large-scale distributions and spread.[5] These processes are interactive, so that local- and regional-scale processes are not independent.[6–8] Ultimately the large-scale dynamics of plants depend on the interplay of processes across scales.[4,9,10] Unravelling this complexity is vital in managing the spread of pest plants such as arable and invasive weeds to enable proactive action against losses to crop yields, and in identifying areas likely to be susceptible in the future.

The primary focus of weed management is usually at the field scale, with individual interventions decided on a field-by-field basis. However, the factors driving weed abundance can frequently be operating at larger scales.[11] These can include changes in land management,[12] climate[13] or evolved herbicide resistance.[14] For such reasons, it is important to understand changes in distributions of weeds beyond the field scale. A great

---

* *Correspondence to: RP Freckleton, Department of Animal & Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK. E-mail: r.freckleton@sheffield.ac.uk*

a *School of Animal Rural & Environmental Sciences, Nottingham Trent University, Nottingham, UK*

b *Department of Animal & Plant Sciences, University of Sheffield, Sheffield, UK*

c *Centre for Ecology and Hydrology, Wallingford, UK*

d *Botanical Society of Britain and Ireland, Harrogate, UK*

deal of insight can potentially be gained from large-scale surveys that link distributions to management practices.[15,16]

Several studies have mapped weeds across large numbers of sites[15,16] and in some countries systematic periodic surveys are undertaken to map changes in distributions.[17] However large-scale atlas surveys require considerable coordination, effort and time. In mapping weeds it also matters whether a population is comprised of just a few plants (i.e. early stages of an infestation) *versus* an extensive infestation of high densities (i.e. a well-established population). For this reason, density-structured modelling approaches were developed with populations being assigned to one of a number of ordinal categories.[18–20] These techniques, however, do not address the problem of rapid mapping at larger spatial scales. Therefore, there is a need to develop methods to rapidly map densities of weeds at scales beyond individual fields, and in this paper we develop a new approach.

Here we analyse the large-scale dynamics of an annual grass, *Alopecurus myosuroides* in the UK. This weed of cereal crops (primarily winter wheat) has become the most significant problem in UK agriculture, costing as much as £400 m per year,[21] and is present in 80% of fields across the arable growing region of the UK.[14] Repeated use of the same chemicals has led to the wide-scale evolution of resistance to herbicides in this and other species,[22–25] and is correlated with high local densities.[14]

The large-scale environmental controls of the geographic distribution of *A. myosuroides* are not well understood, although local-scale determinants of the abundance of this weed are becoming better characterised. Based on large-scale survey data, Hicks *et al.*, for example, describe a strong latitudinal trend in the density of *A. myosuroides*, this being higher in the south of the UK than in the north. Previous coarse-scale atlas surveys have similarly reported latitudinal differences.[26,27] One hypothesis is that the distribution of *A. myosuroides* may be limited by climate: large-scale population modelling suggests that the northward spread of *A. myosuroides* is limited by climate-driven demography.[13] The model in this study showed close correspondence with distribution data collected up to the year 2000, and it was predicted that the current distribution was unlikely to change a great deal even under the most extreme climate change scenarios. Published nearly a decade ago, we can now begin to test these predictions.

It has been suggested that edaphic factors may limit the distribution of *A. myosuroides*. As long ago as the 1920s, it was reported, for instance, that *A. myosuroides* prefers 'heavy land'.[28] Recently, it has been shown that local distributions of *A. myosuroides* are correlated with variations in soil characteristics such as moisture and soil organic carbon.[29,30] However, it is not clear how these local relationships influence variation in density at larger scales.

In this paper, we address two sets of objectives. The first set relates to developing a methodology for characterizing the large-scale distribution of annual weeds, together with variation in density. In particular, (i) we present a rapid method for measuring weed population densities across a landscape and (ii) we show that this method captures the distribution of the species, by comparison with independent atlas data. The second set of objectives are concerned with understanding the distribution of *A. myosuroides*. Specifically, (i) we characterise the large-scale distribution of *A. myosuroides* and map changes in distributions in recent decades, (ii) we relate the occurrence of the species to edaphic and land-use data to test hypotheses about the factors driving abundance (specifically the roles of soil type and climate) and (iii) we use this information to assess the risk of further spread into areas where *A. myosuroides* is currently not a problem.

2 MATERIALS AND METHODS

2.1 UK atlas data

The distributions of plants in the UK have been monitored extensively and published in atlases.[26,27] Here we analyse the data on *A. myosuroides*. Data collection is coordinated by the Botanical Society of Britain and Ireland (BSBI). The first atlas was published in 1962 and collated historical records up until that time.[31] This atlas has been updated in 2002[32] and 2020 (Fig. 1(a)). Here we refer to these as the first, second and third atlases, respectively. Species occurrence are recorded in presence format at the scale of hectads (10 × 10 km).

2.2 Analysis of UK atlas data

The presence within each 10 × 10 km site at the time of the first census (t) was used to predict presence in the second census ($t + 1$). We assumed a binomial error distribution with a logit link. There are known biases in using such data to monitor distribution change.[33,34] We use a simple measure of observation effort[5] (see also Hill[34]). The change in number of species found at each site between censuses (Δ) was included in every statistical model to account for variations in observation effort.[5]

Following Doxford & Freckleton[5] we fitted four statistical models with the aim of characterising the change in distribution of the species over recent decades. The four statistical models make different assumptions about how the distribution of the species might spread. The first is referred to as the *random colonisation* model and assumes that colonisation occurs at random with no spatial effects:

$$\operatorname{logit}\left(p_{i,t+1}\right)=\beta_0+\beta_P P_{i,t}+\beta_\Delta \Delta_i \text{ Model I} \tag{1}$$

$p_{i,t+1}$ is the probability (between 0 and 1) of site i being occupied at time $t + 1$. The intercept β_0 is the probability that an isolated site will be occupied at time $t + 1$. $P_{i,t}$ is the (binary) presence or absence at time t and allows for differential occupancy or extinction depending on whether the site is currently occupied or not and β_P models this dependency. β_Δ is the parameter that controls for variations in observation error (Δ).[5]

Model II is termed the *localised phalanx* model and is developed from Model I by adding an additional term which measures the effect of neighbouring sites (considered as a square lattice, in which each square has eight neighbours, i.e. a Moore neighbourhood) on occupancy in time $t + 1$:

$$\operatorname{logit}\left(p_{i,t+1}\right)=\beta_0+\beta_P P_{i,t}+\beta_N N_{i,t}+\beta_\Delta \Delta_i \text{ Model II} \tag{2}$$

β_N is the strength of the effect of $N_{i,t}$ on occupancy, with $N_{i,t}$ varying between zero and eight populated neighbour sites surrounding focal site i. Increasing $N_{i,t}$ will, according to this model, increase the probability of a site being colonised. However, $N_{i,t}$ only contributes expansion at the periphery of the current distribution.

Model III is the *phalanx-spread* model whereby a distribution has undergone a wide-ranging expansion:

$$\operatorname{logit}\left(p_{i,t+1}\right)=\beta_0+\beta_P P_{i,t}+\beta_{N1} N_{i,t+1}+\beta_\Delta \Delta_i \text{ Model III} \tag{3}$$

Spatial dependency is modelled through the occupancy at time $t + 1$, rather than occupancy at time t: this is an autocovariate in

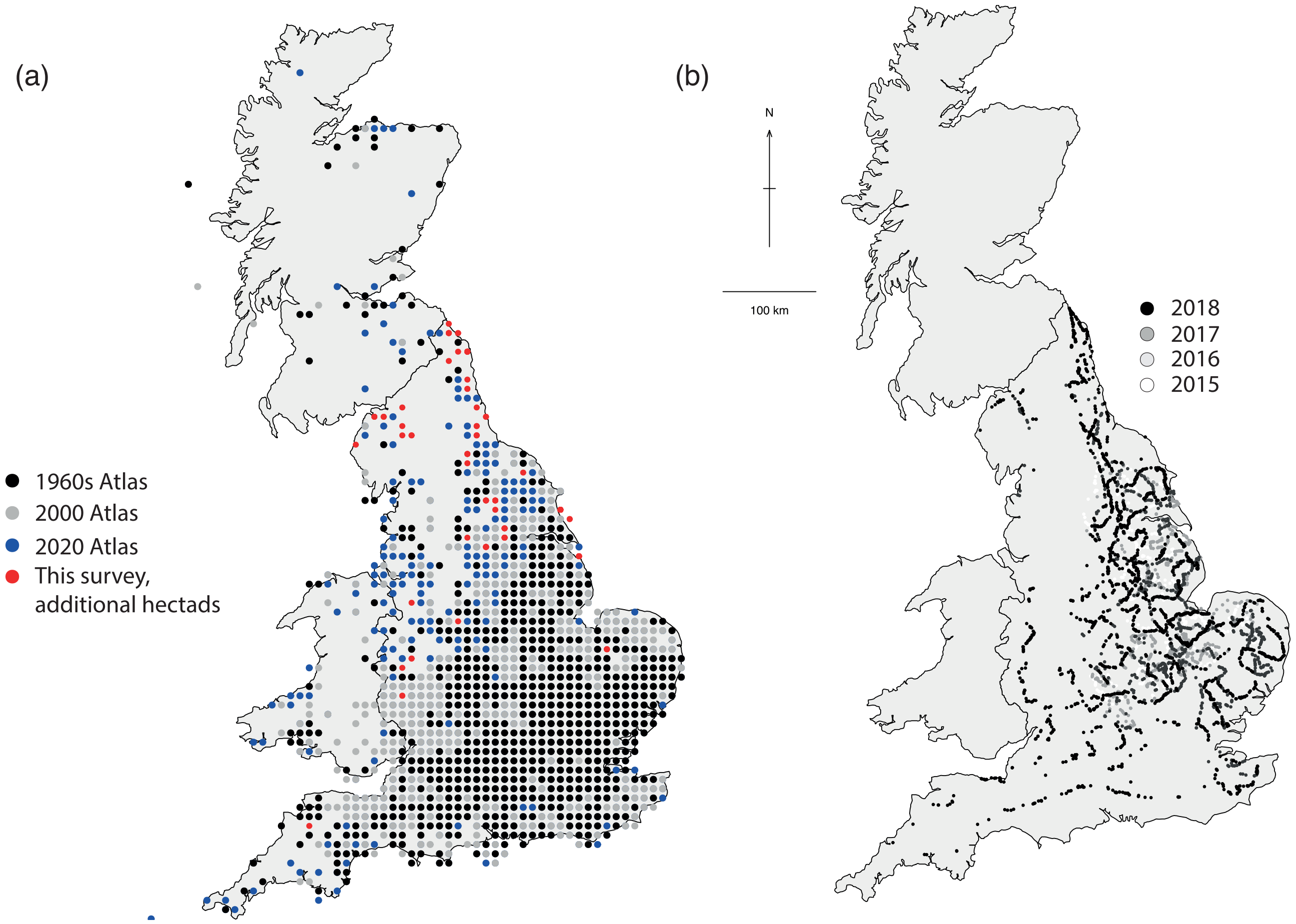

Figure 1. Mapping distributions of *Alopecurus myosuroides*. (a) The distribution of *A. myosuroides* in BSBI atlases and new sites found in this study. (b) The sites surveyed in this study.

which the state of the nearest neighbour sites is used as a predictor of response at site *i*. If a species is spreading (or retreating) as a spatial wave across the landscape, the probability that site *i* will become occupied (or extinct) will be greater in areas in which colonisation (or extinction) have also happened. β_{N1} is the strength of the effect of $N_{i,t+1}$ on occupancy.

$N_{i,t}$ and $N_{i,t+1}$ were also modelled as covariates in the same model, combining the *localised phalanx* and *phalanx-spread models*:

$$\mathrm{logit}\left(p_{i,t+1}\right) = \beta_0 + \beta_P P_{i,t} + \beta_N N_{i,t} + \beta_{N1} N_{i,t+1} + \beta_\Delta \Delta_i \ \text{Model IV} \quad (4)$$

Model IV allows for the initial and eventual configurations of occupied patches to affect occupancy.

We fitted these models to the data in three ways. First, we modelled the change between Atlas 1 and Atlas 2, within which period there was a large expansion of the distribution of *A. myosuroides*. Second, we modelled the change between Atlas 2 and Atlas 3. Finally, we modelled the whole dataset, adding an additional intercept term to measure the difference between the two transition periods (period 1, i.e., Atlas 1 → Atlas 2, and period 2, i.e., Atlas 2 → Atlas 3). We used the Bayes Information Criterion (BIC) to compare the relative fit of the four models.[5]

2.3 Field-scale density assessments

We surveyed densities of *A. myosuroides* within fields across the UK to map large-scale trends in density. We undertook our surveys in 2015, 2016, 2017 and 2018, concentrating on the arable belt of the UK, as well as encompassing the core as well as northern and southern edges of the distribution. The skeleton for the sampling program was a system of 70 farms established for a detailed farm-level study of the population dynamics of *A. myosuroides*.[14] Observers drove along local roads and recorded the density of *A. myosuroides* in winter wheat fields. The team consisted of an observer accompanied by a driver who, for safety reasons, had no responsibility for making observations. When an arable field was encountered, the car stopped and the observer recorded the location and estimated the population of *A. myosuroides* (see below). In 2017 we extended the survey region to the north-west (Cumbria), to the south-east (Kent) and the south-west (Oxfordshire) to permit us to more fully capture clines in weed abundance. In 2018 the survey was repeated and extended to cover the entire arable area of England.

The location, crop type (filtered to winter wheat) and an estimate of field-scale density were recorded from the side of each field. The rapid assessments used a five-point density structured approach[14]: *absent* is a field containing no weeds, while four states from *low* to *very high* represent increasing population densities. Visual representations of each of the latter four density states are given in Supplementary Online Material (SOM) Table 1.

The measurements of density are treated as point locations. The datasets on environmental data (below) are gridded data (1 km^2 squares). For each field we determined the square in which it was located and recorded the mean value of the respective variable.

Table 1. Models for the spread of *Alopecurus myosuroides* inferred from atlas data (Fig. 1(a)) following Doxford & Freckleton[5]

(a)	Δ	(±se)	*P*(*t*)	(±se)	*N*(*t*)	(±se)	*N*(*t* + 1)	(±se)	BIC	df
Model I	0.01	0.00	4.35	0.17					783.38	3
Model II	0.01	0.00	2.82	0.20	0.73	0.04			337.90	4
Model III	0.01	0.00	2.97	0.21			0.71	0.03	0.00	4
Model IV	0.01	0.00	2.98	0.22	−0.02	0.06	0.72	0.05	7.77	5

(b)	Δ	(±se)	*P*(*t*)	(±se)	*N*(*t*)	(±se)	*N*(*t* + 1)	(±se)	BIC	df
Model I	0.00	0.00	3.44	0.12					642.25	3
Model II	0.00	0.00	1.21	0.17	0.54	0.03			252.10	4
Model III	0.00	0.00	1.25	0.16			0.64	0.03	0.00	4
Model IV	0.00	0.00	1.20	0.19	0.03	0.05	0.63	0.04	7.60	5

(c)	Δ	(±se)	*P*(*t*)	(±se)	*N*(*t*)	(±se)	*N*(*t* + 1)	(±se)	BIC	df
Model I	0.01	0.00	3.65	0.09					1463.81	4
Model II	0.01	0.00	1.73	0.12	0.57	0.02			682.26	5
Model III	0.00	0.00	1.80	0.12			0.66	0.02	0.00	5
Model IV	0.00	0.00	1.91	0.13	−0.07	0.04	0.71	0.03	5.08	6

We fit four models: Model I, the random colonisation model; Model II, the localised phalanx model; Model III, the phalanx-spread model; and Model IV that integrates elements of all models. The tables report the parameter estimates for the effect of a proxy of sampling effort (Δ), the previous state of each hectad (*P*(t)) and the numbers of neighbours containing *A. mysosuroides* (*N*) at times *t* and *t* + 1. The Bayes Information Criterion (BIC, relative to the best model) and overall model degrees of freedom. (a) Models for change in distribution between atlas 1 and atlas 2. (b) Models for change in distribution between atlas 2 and atlas 3. (c) Models for change in distribution over both atlas periods, including an additional intercept term measuring the difference between the atlas periods.

2.4 Soil data

Data on soil types were taken from LandIS (https://www.landis.org.uk/services/soilsguide/series_list.cfm). This system recognises 753 soil series, which may be simplified to 11 major soil types.[35] Of the 11 possible soil types, our surveys included significant numbers of sites with four soil types: Type 4 (pelosols), Type 5 (brown soils), Type 7 (surface-water gley soils) and Type 8 (ground-water gley soils), the frequencies of which we included as predictors in our statistical models. Of these soil types, the gleys are 'heavy' soil types, associated with seasonal water logging. Pelosols have a high proportion of clay, which tends to dry out. Brown soils represent a diverse range of soils commonly in agricultural use. The resolution of the UK soil map is 1 × 1 km. For each monitored location we measured the proportion of each of these soil types within the 1 km^2 square in which it is situated.

2.5 Rainfall data

We used the CEH-Gridded Estimates of Areal Rainfall (CEH-GEAR) dataset which consists of 1 × 1 km gridded estimates of daily and monthly rainfall for Great Britain and Northern Ireland.[36,37] This is a dataset of modelled rainfall that extends up to 2015 at the time of analysis. We used data on mean monthly rainfall. Data on annual rainfall are not available for each of the individual years in which we monitored *A. myosuroides*. We were therefore unable to analyse whether interannual variation in rainfall drives year-to-year changes in abundance. Consequently, we analysed whether the abundance of *A. myosuroides* varies with average annual rainfall. The rainfall data did not exactly match our survey period, therefore we used mean rainfall from the four most recent years for which data were available (2012–2015). The rainfall variable we used therefore measures the degree to which longer-term patterns of rainfall influence current densities. We did not use longer averages (e.g. 30 years) because we wished to test whether recent rainfall has influenced densities. We averaged the mean monthly rainfall across these 4 years for the 1 × 1 km square within which each monitored location is situated.

2.6 Land use data

Information on prevailing patterns of land use were taken from the CEH Land Cover Map.[1] Given that our objective is to map the occurrence of *A. myosuroides* within arable landscapes, for each locality, we recorded this proportion for the 1 km^2 square within which it was located.

2.7 Analysis of field-scale density data

We modelled our field-scale density data using generalized additive models (GAMs),[38] which are a family of flexible modelling approaches. This approach is appropriate for this problem because it permits us to model large-scale geographical clines in density through flexible smoothing functions, as well as the effects of individual drivers (soil and rainfall) included as conventional regression covariates. Response and predictor variables were normalised prior to fitting.

The starting point is an ordinary linear model. $N_i(t)$ is the density state at location *i*. This is modelled as a function of $j = 1\ldots4$ soil variables measured for site *i* (S_{ij}) and rainfall average at each location (R_{ij}). Initial analysis indicated that the ordinal response variable *N* (numerically coded 0…4) could be treated as continuous following a square root transformation (denoted $n = N^{0.5}$) and that this adequately satisfied the distributional assumptions of the model (we also compared our results with those of models in which the raw counts, *N*, were assumed to follow an ordered categorical distribution). The fitted models and outcomes in terms of

---

[1] <http://digimap.edina.ac.uk>, downloaded: 2017-2110-18 20:50:30.58.

significance tests were nearly identical, but the interpretation of the model using the continuous response is simpler. The model is then:

$$n_{i,t}=\alpha_t+\sum_{j=1}^{4}\beta_j S_{ij}+\beta_R R_i+\varepsilon_{i,t} \tag{5}$$

where α_t is the year-specific intercept, β_j is the slope measuring the effect of the *j*th soil variable and β_R is the slope for the effect of rainfall. To account for systematic spatial variation in the distribution, we included longitude and latitude (denoted X_i and Y_i, respectively) measured as easting and northing on the national grid. These were included within a thin plate regression spline smooth term [$\tau(X_i, Y_i)$] in the GAM:[39]

$$n_{i,t}=\alpha_t+\sum_{j=1}^{4}\beta_j S_{ij}+\beta_R R_i+\tau(X_i,Y_i)+\varepsilon_{i,t} \tag{6}$$

Analysis of this model indicated that there was significant spatial positive autocorrelation (see below for details of testing), that is, residuals from locations that are close together tended to be more similar. This could be because of the influence of local geographic factors that we were unable to measure or include in our models, or because the predictors are averaged at a relative coarse scale (1 km^2 square) relative to the observations, with potentially several measurements within the same square or very close together. Such autocorrelation can significantly detract from the performance of such models.[40] We therefore added a grouping representing the hectad (10 × 10 km^2) within which each observation was recorded.

$$n_{i,t}=\alpha_t+\sum_{j=1}^{5}\beta_j S_{ij}+\beta_R R_i+\tau(X_i,Y_i)+h(X_i,Y_i)+\varepsilon_{i,t} \tag{7}$$

$$h(X_i,Y_i)\sim N\left(0,\sigma_h^2\right)$$

$h(X, Y)$ represents the hectad that contains point X, Y. This is modelled as a random effect with variance σ_h^2. By grouping observations from the same hectads in this way, it effectively acts as a local-level random effect, which removes variance owing to locality. Although this approach is slightly crude (e.g. it does not account for covariance between adjacent hectads), we tested the success of the method by retesting the residuals for spatial autocorrelation. The models were fitted using the gam() function in the MGCV package in R.[41] We used restricted maximum likelihood to select the model as this has been shown to be more robust to undersmoothing, although it is computationally slightly more intensive.[31]

As noted above, we tested for spatial autocorrelation in the model residuals. We used Moran's I to test for this using a randomisation approach,[32] implemented in the ade4 library[42]). According to this test, Eqn. (7) yielded a fit in which there was no spatial autocorrelation in the residuals of the final model.

2.8 Model analysis

The fitted GAM was analysed in three ways. First, to visualise the geographical variation in field densities of *A. myosuroides* across the study area, we plotted the fitted smooth term, that is, excluding the fixed effects as a time-varying intercept model:

$$\hat{n}_{i,t}=\hat{\alpha}_t+\hat{\tau}(X_i,Y_i) \tag{8}$$

Second, to visualise the full variation in density resulting from both broad-scale geographic trends and the influence of the fixed predictor variables we generated the full fitted model:

$$\hat{n}_{i,t}=\hat{\alpha}_t+\sum_{j=1}^{5}\hat{\beta}_j S_{ij}+\hat{\beta}_R R_i+\hat{\tau}(X_i,Y_i) \tag{9}$$

Thirdly, to visualise the influence of the fixed effect predictors we mapped:

$$\hat{n}_{i,t}=\hat{\alpha}_t+\sum_{j=1}^{5}\hat{\beta}_j S_{ij}+\hat{\beta}_R R_i \tag{10}$$

This final quantity is important as it measures the influence of environmental variation, controlling for large-scale geographic variation. Current variation in density bears an imprint of historical land management and previous densities of the species. By eliminating these, Eqn. (10) is a measure of the *potential* density of *A. myosuroides* in the absence of such confounding or historical factors, that is, a measure of risk of infestation. We normalised this quantity (scaled to a value between 0 and 1).

2.9 Comparison of independent datasets

To tie the two datasets together, we compared the predictions of the GAM with the atlas data. Note that the model predictions and atlas data are *independent* of each other because the latter were not used in the construction of the GAM. We used Mann–Whitney–Wilcoxon tests (mathematically equivalent to the area under the curve from a receiver operator characteristic plot) to compare the predicted densities from the GAMs with the atlas distribution of *A. myosuroides*. We compared the predictions of the fixed effects (continuous values predicted from Eqn. (10)) for England and Wales with recorded atlas distributions (binary presence/absence) and calculated the rank biserial correlation (r_{sb}; this varies between −1 and 1) as a measure of effect size.

3 RESULTS

3.1 Large-scale distribution and change in atlas data

Based on atlas data, the distribution of *A. myosuroides* is primarily southerly, but appears to have shown a steady northerly increase over the past 50 years (Fig. 1(a)). The distribution of the species has increased (from 655 hectads in the 1960s atlas, to 948 in the 2000 atlas and occupies 963 hectads in the latest). There is low but increasing frequency in the north of England, with very low occurrence in Scotland.

Modelling the change in distribution between atlas surveys our results confirm that the species has indeed spread, in contrast to previous predictions.[13] We found that the best fitting model was Model III, the 'phalanx spread' (Table 1), and this was true for all time periods. This model includes an autoregressive term for the number of neighbours colonised surrounding a focal hectad, indicating that spread is contiguous and relatively localised.

3.2 Spatio-temporal variation in field-scale density assessments

Overall, our field surveys showed that *A. myosuroides* is extremely common, occurring in 65% of surveyed fields. This fraction was lowest in 2018 (51%), but highest in 2016, with all surveyed fields containing *A. myosuroides*. The modal density state was the 'low' density state, with 44% of fields determined to be in this state. There is variation in weed density both geographically as well as

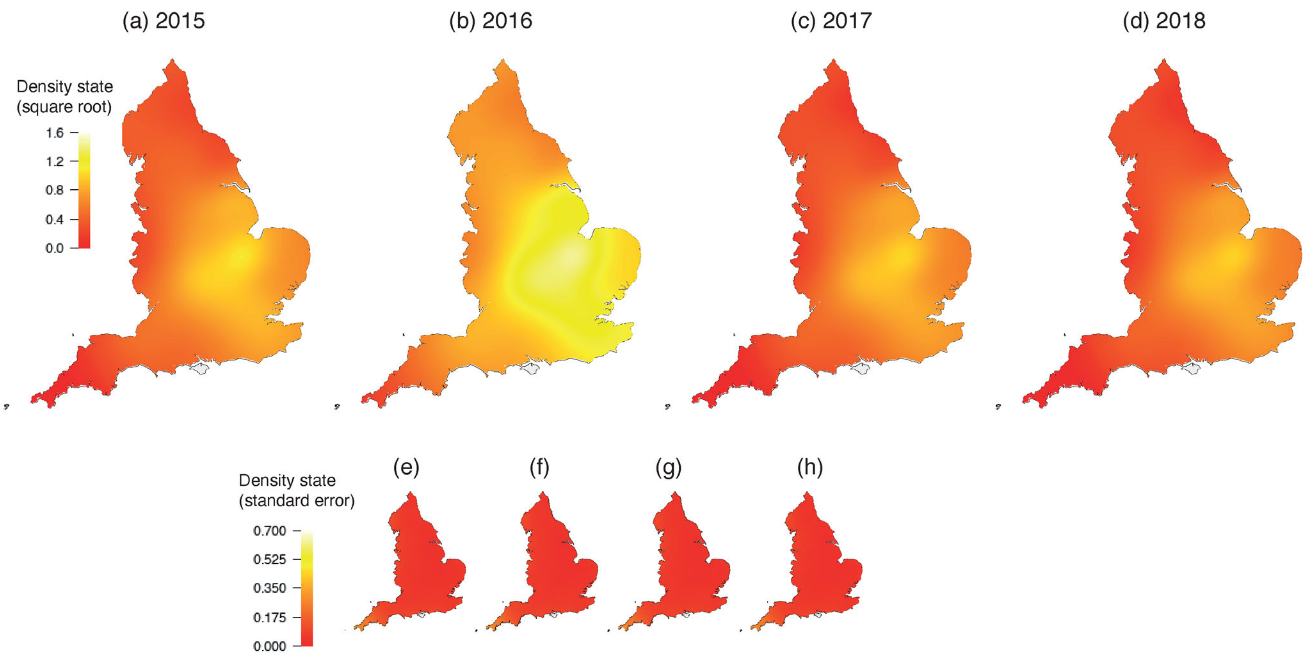

Figure 2. Modelled density of *Alopecurus myosuroides* in (a) 2015, (b) 2016, (c) 2017 and (d) 2018. The maps show generalized additive models to the survey data, describing geographical variation through a thin plate regression spline (see text for details). The standard errors (units of density) of the fitted values are shown in (e)–(h).

between years: according to the GAM, Fig. 2 shows the fitted densities in each of the years separately (Eqn. (8)). The highest densities are observed in the central area of England, along an approximate belt running from the south of our study area, through the centre. Densities are notably lower in the northern extreme, as well as in the far south east and west.

3.3 Drivers of occurrence in in field-scale density assessments

Generalized additive models fitted to the national field-scale density assessments indicated a role for both soil type and environmental conditions in driving densities, however there was considerable interannual variation that is not explained by average conditions (Table 2). Of the soil types considered there were clear positive correlations of density with the occurrence of both types of gley soils (Table 2). There was a negative impact of rainfall on the mean density of *A. myosuroides* (Table 2) with a decrease in densities with increasing mean rainfall. Although in the UK the areas of highest rainfall are in the uplands, our surveys excluded such areas. Thus, the impacts of average rainfall on density are through effects that occur within the range of conditions experienced in arable farmed landscapes.

Table 2. Generalized additive models (GAMs) for data from the field-scale density assessments of *A. myosuroides* within the surveyed area (see Fig. 1(b))

Term	Coefficient	SE	df	*F*	*P* value
Year intercepts	0.75	0.085	3	155.1	$<1 \times 10^{-6}$
+2016	0.59	0.046			
+2017	−0.13	0.047			
+2018	−0.19	0.043			
Soil type					
Pelosol	0.045	0.023	1	3.71	0.054
Brown soil	−0.038	0.023	1	2.60	0.107
Surface-water gley	0.051	0.023	1	4.82	0.028
Ground-water gley	0.090	0.025	1	13.1	< 0.001
Weather					
log(rainfall)	−0.202	0.020	1	100.4	$<1 \times 10^{-6}$
Smooth term					
$t(X, Y)$	-	-	0.99	618.4	$<1 \times 10^{-6}$

Separate intercepts were fitted for each year. Four soil types were abundant in the area surveyed, and the proportion of the hectad's area were included as linear covariates. Mean rainfall (2012–2015) was included, given previous suggestions that moisture plays a role in driving the abundance of *A. myosuroides*.

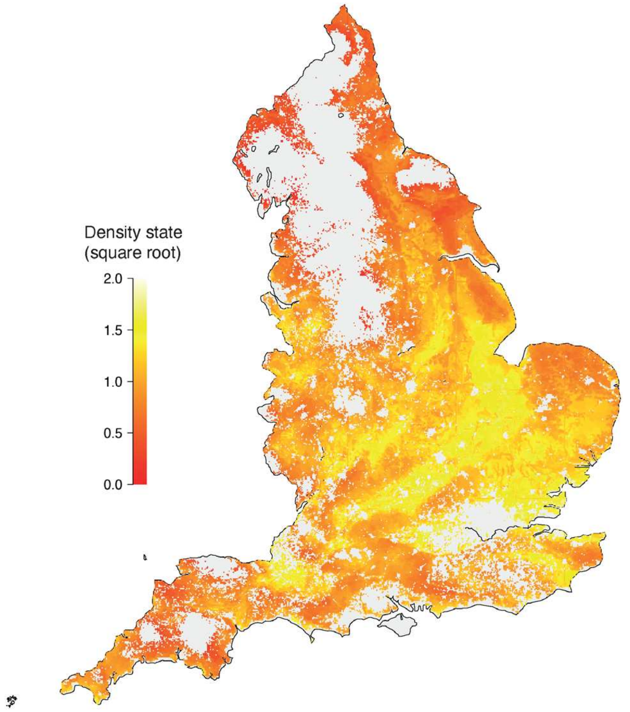

Figure 3. Map of modelled density of *Alopecurus myosuroides*. The model includes soil type and rainfall (Table 2) as well as a smooth term (thin plate regression spline, see text for details).

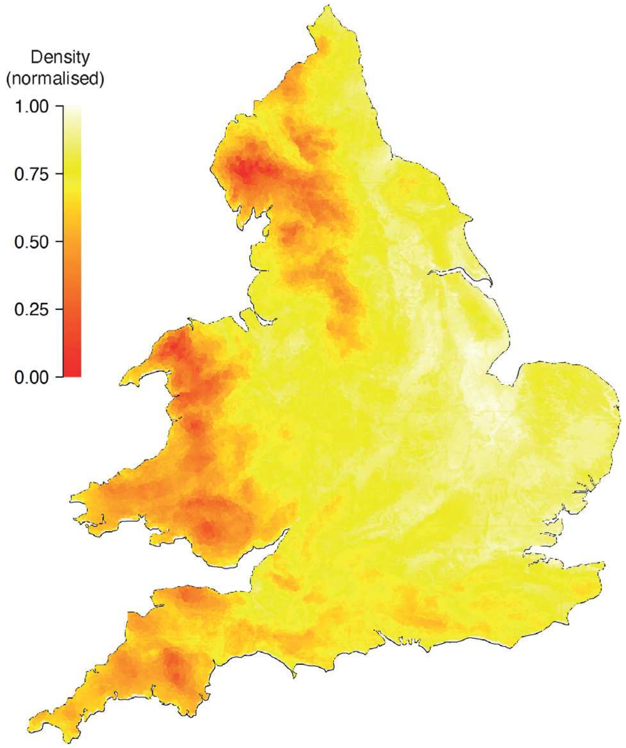

Figure 4. Risk modelling of *Alopecurus myosuroides*: map of modelled density of the whole of England and Wales. The smooth term has been removed so the map shows variation in density attributable to soil and rainfall, that is, relative suitability in terms of edaphic conditions only.

3.4 Predicted map of *A. myosuroides* density

Based on the data from the field-scale density assessments Fig. 3 represents the map of current average density of *A. myosuroides* in UK arable agriculture predicted from the full GAM (Eqn. (9)). Reflecting the primarily arable focus of our surveys, in Fig. 3 all areas with less than 5% arable land use cover have been masked out. The mapped density is the net consequence of both the systematic longitudinal and latitudinal variation in densities, together with the modelled effects of environmental drivers within the surveyed area.

Although we have identified significant environmental predictors of densities, the latitudinal cline of density is very clear in Figs 2 and 3. These large-scale clines will result, in some part, from historical factors. To eliminate these effects, in Fig. 4 the latitudinal and longitudinal trends have been removed. This indicates that historical factors are likely to have played a significant role in driving density. None of the variables identified (rainfall or soil type) show such an obvious latitudinal trend.

3.5 Predicting distributions

In Fig. 4 we have extrapolated the GAM model to the whole of England and Wales, controlling for large-scale variation (Eqn. (10)). This map therefore is a measure of potential habitat suitability. Comparing Fig. 4 (based on the model for our field-scale density assessments) with Fig. 1 (the map of current occurrence, based on the UK atlases), it appears that the areas of occupancy (points in Fig. 1(a)) coincide with the areas of predicted highest densities (yellow shaded areas in Fig. 4). Note that these two maps are independent of each other because the data from Fig. 1(a) do not enter the models used to generate Fig. 4.

To synthesise the results from the two datasets, we compared the observations from the atlas data with the model predictions. Accordingly, Fig. 5 shows a more formal comparison of the two datasets: there is clearly good discrimination between the occupied and unoccupied atlas hectads. The correspondence between the two datasets has increased over the past 50 years, indicating models based on current densities predict current occupancy more accurately that historical occupancy: the biserial correlation for the correspondence between model predictions and the occupancy in the 1960s atlas is 0.44, but this increases to 0.64 for current occupancy, including the data from this study.

4 DISCUSSION

At a national scale the UK, in common with many other countries, lacks schemes for regular large-scale mapping of weeds. Large-scale atlas data are a rich source of information on long-term changes[17] but are undertaken too infrequently to allow rapid changes in species' distributions to be mapped that occur as a consequence of changes such as rapid evolution of herbicide resistance,[15] land use change[43] or dispersal.[44] Here we have shown how a large-scale sampling approach can be used to accurate map occurrence, predict susceptible areas and test long-standing hypotheses about possible drivers. Our own data showed, for example, that *A. myosuroides* has spread recently, so that the most recent atlas data are already in need of updating as the weed is now occurring in areas in which it was previously not reported. This also highlights the need to regularly survey outside the current areas of occupancy to detect new populations.

4.1 Management of *A. myosuroides*: spatial and temporal perspective

Our models show that the spread of the weed has occurred through a local phalanx. This implies that local processes drive

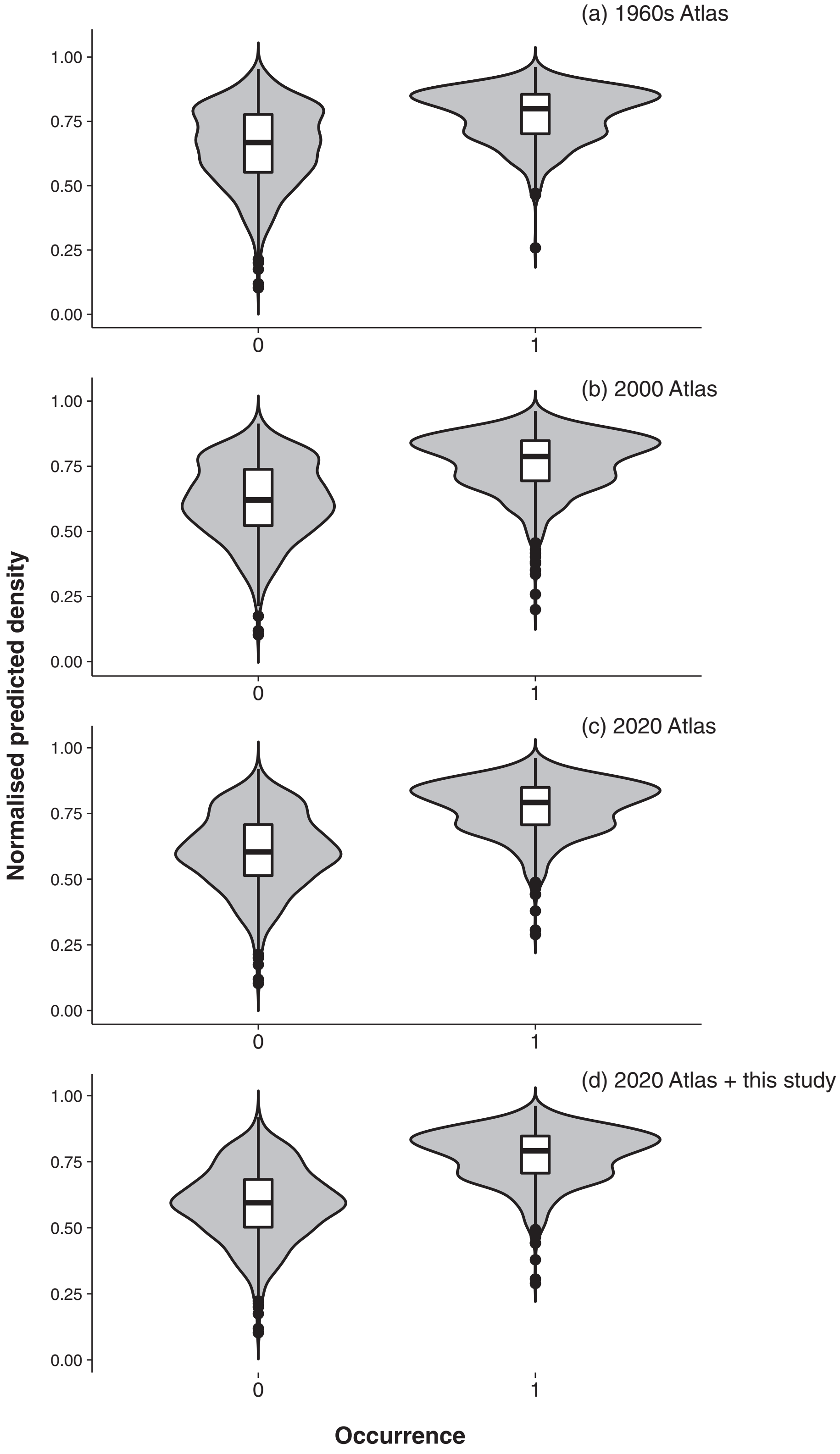

Figure 5. Comparison of predicted densities based on model predictions based on the national field-scale density assessments (Fig. 4) with occupancy recorded in the atlas data (Fig. 1(a)). In each case we have compared the predictions from the model with the records of presence and absence from one of the atlases: (a) 1960s atlas (Wilcoxon test $r_{rb} = 0.44$, $P < 0.0001$); (b) 2000 atlas (Wilcoxon test $r_{rb} = 0.52$, $P < 0.0001$); (c) 2020 atlas (Wilcoxon test $r_{rb} = 0.62$, $P < 0.0001$); (d) 2020 atlas supplemented with the data collected in this study (Wilcoxon test $r_{rb} = 0.64$, $P < 0.00001$). Note the *y* axis presents normalised density, relative to minimum and maximum predicted densities, hence is on a 0–1 scale.

the spread of this weed. Four local-scale processes are important. First, increases in local densities have been driven by herbicide resistance,[14] and local spread will be facilitated by the movement of seeds and pollen with resistant genotypes. Second, seeds may disperse through natural dispersal, although *A. myosuroides* has no specialised mechanisms for dispersal beyond a few metres,[45,46] indicating a limited role for this mechanism in geographic spread. Third, seed can be moved on farm machinery: there has been no formal analysis of this form of dispersal in *A. myosuroides*, but this is a well-known mechanism of seed dispersal in agricultural systems.[44] Although there is some industry recommendation that cleaning of machinery is important,[47] the results here indicate that spread of the weed into new areas has been considerable.

Over the 4 years, densities were substantially higher in 2016 than in the three other years (Fig. 2). The decline in weed abundance in 2017 compared with 2016 mirrored a national-scale 1.9% decrease in winter wheat area.[2] This decline in winter wheat frequency is attributable to the difficulty of producing wheat in fields with high infestations of *A. myosuroides*, which often prompts farmers to grow alternative cereals such as spring barley.[48] The reason for the increase in 2016 is not fully explained: it seems likely that weather would play a role, as this often impacts on the effectiveness of control. Future analyses would benefit from finer-scale weather information to test this possibility.

4.2 The importance of edaphic factors

It is important to understand the impacts of environmental conditions on both the fine- and large-scale distributions of weeds. Such understanding may permit us to forecast future distributions.[29] At the fine scale associations have been found between soil properties and the occurrence of *A. myosuroides*.[30] This includes positive correlations between soil texture (relating to clay content) and weed density, as well as a positive effect of soil moisture. These relationships within-field, that is within local populations,[30] exhibit scale-dependence, and in this paper we have analysed these correlations at larger scales.

We found a positive association between the occurrence of *A. myosuroides* and heavier gley soils at the landscape scale. The existence of this relationship confirms previous suggestions that *A. myosuroides* is associated with ‘heavy’ land,[28] although earlier work[49] suggested that prevailing land use may be more significant in driving abundance.[50] The negative effect of rainfall on the abundance of *A. myosuroides* that we report may well reflect this: in the UK rainfall is correlated with altitude, with higher rainfalls at greater elevations. In Fig. 4, the areas of predicted low densities (red areas) correspond to the uplands. Upland areas are unfavourable for arable farming, and hence are likely to be unsuitable for *A. myosuroides*. Indeed, this association with lowlands is evident in the overall distribution (Fig. 1(a)).

4.3 Large-scale distribution

There are clear geographic trends (Figs 2 and 3) in the current distribution, with a ‘corridor’ of high densities through the central portion of England, one of the highest intensity arable regions of the UK. Comparing the current modelled distribution of densities (Fig. 3) with the potential modelled distribution of densities in terms of soil and rainfall (Fig. 4), there are many areas in which *A. myosuroides* is currently found at low densities, but that do not appear any less suitable in terms of potential distribution. This map therefore indicates that there are significant fractions of the country which are suitable, but in which infestations of *A. myosuroides* are currently low or nonexistent. In Fig. 5 the predicted normalised density for unoccupied sites is clearly greater than zero, reflecting this.

Previously, the (then) current and future distributions of *A. myosuroides* were forecast using a process-based model.[13] Based on detailed submodels for plant growth and population dynamics, the model forecast future distributions under different climate change scenarios. The model suggested that *A. myosuroides* would be unlikely to spread more than modestly outside of the current distribution, with population growth rates forecast to be negative in most areas unoccupied in the 2000 atlas survey (i.e. outside the black and grey points in Fig. 1). Clearly the data recorded both in the 2020 atlas and our surveys (blue and red points in Fig. 1) indicate that these predictions have not been borne out, with our fitted models showing that the species has been spreading consistently over all periods.

Indeed, our models indicate that the species is likely to spread further, with much of the mismatch between Fig. 3 and Fig. 4 likely to be the consequence of slow spread, with the species yet to fill all potential habitat. The likely reasons for the continued spread of the species could include herbicide resistance.[14] Increases in abundance owing to this generate large source populations from which the species is able to spread (e.g. through machinery). In the light of this, managing uninvaded areas to minimise development of herbicide resistance should be a priority for management, for example through increased cultural management.

Notwithstanding large-scale trends in density, the map of predicted suitability (Fig. 4) provided a good predictor of current occupancy (Fig. 5). The increase in correspondence between the maps (i.e. increase in bi-serial correlation) from pre-1960s to the modelled distribution is a result of the expansion of distribution of *A. myosuroides*. Populations have expanded into areas which were previously unoccupied, but clearly suitable in terms of soil, rainfall and land use. The species therefore has historically underfilled its potential niche. Based on this finding, there are good reasons to expect that *A. myosuroides* could expand its range into currently unoccupied areas which Fig. 4 indicates may be suitable for the species.

4.4 Methodology

In monitoring populations there is a trade-off between the detail in which populations can be measured and the spatial scale over which sampling can be extended, which depends on the resources available.[51] Our approach is simpler than conventional, time-consuming plant count survey techniques. It is easy to train surveyors with high levels of repeatability and the resultant data can be used to answer questions that are difficult to address using data collected only at the scale of local populations. The rapid nature of the technique also allows surveys to cover large geographical areas (either within the field or between fields) and to answer questions at much larger scales than previous weed research has been able to answer.

In this study *A. myosuroides* is highly visible when mature, and this approach could be modified for a range of weeds that are similarly easy to identify at long range. We have previously developed methods for in-field monitoring that allow us to monitor this and a suite of other highly visible weed species.[18] We have

---

[2] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/670004/structure-jun2017final-uk-21dec17.pdf.

confirmed that for such weeds, rapid visual surveys provide an accurate and repeatable assessment of abundance. These methods will obviously be most applicable to weeds that are taller than the surrounding crop, with clearly visible differences in colour and growth form. Weeds with shorter growth forms, particularly those that are lower than the crop, would be impossible to monitor using such approaches. In contrast *A. myosuroides* is a highly visible weed (see SOM Table 1). Indeed it is so evident that remote mapping is possible, for example using unmanned aerial systems.[52,53]In the future, the logical extension of the method we report here would be a remote autonomous monitoring platform.

A number of studies have monitored weed populations at large spatial scales, and the methodology employed is variable.[15,54,55] The factors driving choice of methods probably relate to available resources in terms of labour and funding, and there is not a current standard methodology for such monitoring. There are, however, potential advantages to standardisation of methods. Creation and use of standardised methods for data collection can aid the collation of data on both spatial and temporal scales, enabling the collection of spatially wide-reaching, long-term datasets. Using a standardised methodology can provide vast improvements in data quality when compared to datasets created from bringing together piecemeal bits of data from different studies using differing methods. It is not unusual for several teams to be working simultaneously on the same weed species. Moreover, there is often a lag from collection to publication, if indeed data are published at all, and many datasets undoubtedly lie unused. The likelihood of reuse is greater if the data have been collected in a standardised way. Moreover, agronomists and farmers regularly walk fields and record data. Standardised measures of weed incidence or abundance should be useful in making crop management recommendations. We therefore argue that current trends towards openness in data archiving to promote reuse are best served if we can use standardised methodologies where possible.

REFERENCES

1 Watkinson AR, Density-dependence in single-species populations of plants. *J Theor Biol* **83**:345–357 (1980).
2 Watkinson AR, Freckleton RP and Forrester L, Population dynamics of *Vulpia ciliata*: regional, patch and local dynamics. *J Ecol* **88**: 1012–1029 (2000).
3 Ellner SP, Snyder RE and Adler PB, How to quantify the temporal storage effect using simulations instead of math. *Ecol Lett* **19**: 1333–1342 (2016).
4 Bullock JM, Brown VK, Peel S, Brown R, Everett S, Gray A *et al.*, Plant dispersal and colonisation processes at local and landscape scales, in *Dispersal Ecology*, ed. by Bullock JM, Kenward R and Hails R. Blackwell Science, Oxford (2002).
5 Doxford SW and Freckleton RP, Changes in the large-scale distribution of plants: extinction, colonisation and the effects of climate. *J Ecol* **100**:519–529 (2012).
6 Holt RD, Lawton JH, Gaston KJ and Blackburn TM, On the relationship between range size and local abundance: back to basics. *Oikos* **78**: 183–190 (1997).
7 Buckley HL and Freckleton RP, Understanding the role of species dynamics in abundance-occupany relationships. *J Ecol* **98**:645–658 (2010).
8 Freckleton RP, Gill JA, Noble D and Watkinson AR, Abundance-occupancy relationships and the scaling from local to regional population size. *J Anim Ecol* **74**:353–264 (2005).
9 Freckleton RP and Watkinson AR, The large scale spatial dynamics of plants: metapopulations, regional ensembles and patchy populations. *J Ecol* 419–434 (2002).
10 Dorken ME, Freckleton RP and Pannell JR, Small-scal and regional spatial dynamics of an annual plant with contrasting sexual systems. *J Ecol* **105**:1044–1057 (2017).
11 Cousens R, Wallinga J and Shaw M, Are the spatial patterns of weeds scale-invariant? *Oikos* **107**:251–264 (2004).
12 Firbank LG, Mortimer AM and Putwain PD, *Bromus sterilis* in winter wheat: a test of a predictive population model. *Aspects Appl Biol* **9**: 59–66 (1985).
13 Stratonovitch P, Storkey J and Semenov MA, A process-based approach to modelling impacts of climate change on the damage niche of an agricultural weed. *Glob Chang Biol* **18**:2071–2080 (2012).
14 Hicks HL, Comont D, Coutts SR, Crook L, Hull R, Norris K *et al.*, The factors driving evolved herbicide resistance at a national scale. *Nat Ecol Evolut* **2**:529–536 (2018; in press).
15 Evans JA, Trabel PJ, Hager AG, Schutte B, Wu C, Chatham LA *et al.*, Managing the evolution of herbicide resistance. *Pest Manag Sci* **72**:74–80 (2016).
16 Herrmann J, Hess M, Strek HJ, Richter O and Beffa R, Linkage of the curren ALS-resistance status with field history information of multiple fields infested with black grass. *Julius-Kühn-Archiv* **452**:43–49 (2016).
17 Rich TCG and Woodruff ER, Changes in vascular plant floras of England and Scotland between 1930 - 1960 and 1987 - 1988: the BSBI monitoring scheme. *Biol Conserv* **75**:217–229 (1996).
18 Queenborough SA, Burnet KM, Sutherland WJ, Watkinson AR and Freckleton RP, From meso- to macroscale population dynamics: a new density-structured approach. *Methods Ecol Evol* **2**:289–302 (2011).
19 Freckleton RP, Sutherland WJ, Watkinson AR and Queenborough SA, Density-structured models for plant population dynamics. *Am Nat* **177**:1–17 (2011).
20 Taylor CM and Hastings A, Finding optimal control strategies for invasive species: a density-structured model for *Spartina alterniflora*. *J Appl Ecol* **41**:1049–1057 (2004).
21 Varah A, Ahodo K, Coutts SR, Hicks HL, Comont D, Crook L *et al.*, The costs of human-induced evolution in an agricultural system. *Nat Sustain* **3**:63–71 (2020).
22 Heap IM, The occurrence of herbicide-resistant weeds worldwide. *Pest Sci* **51**:235–243 (1997).
23 Neve P, Norsworthy JK, Smith KL and Zelaya IA, Modelling glyphosate resistance management strategies for palmer Amaranth (*Amaranthus palmeri*) in cotton. *Weed Technol* **25**:335–343 (2011).
24 Delye C, Jasieniuk M and Le Corre V, Deciphering the evolution of herbicide resistance in weeds. *Trends Genet* **29**:649–658 (2013).
25 Neve P, Challenges for herbicide resistance evolution and management: 50 years after Harper. *Weed Res* **47**:365–369 (2007).
26 Perring F, Walters SM. *Atlas of the British Flora: Botanical Society of the British Isles*, London, Thomas Nelson and Sons Ltd; 1962.
27 Preston CD, Pearman DA and Dines TD, *New Atlas of the British and Irish Flora*. Botanical Society of the Britain and Ireland, Oxford, Oxford University Press; (2002).
28 Brenchley WE, *Weeds of Farm Land*. Longman, London (1920).
29 Metcalfe H, Milne AE, Webster R, Lark RM, Murdoch AJ and Storkey J, Designing a sampling scheme to reveal correlations between weeds and soil properties at multiple spatial scales. *Weed Res* **56**:1–13 (2016).
30 Metcalfe H, Milne AE, Webster R, Lark RM, Murdoch AJ, Kanelo L *et al.*, Defining the habitat niche of blackgrass (*Alopecurus myosuroides*) at the field scale. *Weed Res* **58**:165–170 (2018).
31 Wood SN, Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J Roy Stat Soc B* **73**:3–36 (2011).
32 Haining R, *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge University Press, Cambridge (1990).
33 Telfer MG, Preston CD and Rothery P, A general method for measuring relative change in range size from biological atlas data. *Biol Conserv* **107**:99–109 (2002).
34 Hill MO, Local frequency as a key to interpreting species occurrence data when recording effort is not known. *Methods Ecol Evolut* **3**: 195–205 (2012).
35 Avery BW. Soil Classification for England and Wales (Higher Categories). Harpenden: Soil Survey Technical Monograph No. 14.; 1980.
36 Tanguy M, Dixon H, et al. Gridded estimates of daily and monthly areal rainfall for the United Kingdom (1890-2015) [CEH-GEAR]. NERC Environmental Information Data Centre; 2016.
37 Keller VDJ, Tanguy M, Prosdocimi I, Terry JA, Hitt O, Cole SJ *et al.*, CEH-GEAR: 1 km resolution daily and monthly areal rainfall estimates for

the UKfor hydrological and other applications. *Earth Syst Sci Data* **7**: 143–155 (2015).

38 Wood SN, *Generalized Additive Models: An Introduction with R*. Chapman & Hall, London (2016).

39 Wood SN, Thin-plate regression splines. *J Roy Stat Soc B* **65**: 95–114 (2003).

40 Guélat J and Kéry M, Effects of spatial autocorrelation and imperfect detection on species distribution models. *Methods Ecol Evolut*. **9**: 1614–1625 (2018).

41 Team RC, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2017).

42 Dray S and Dufour AB, The ade4 package: implementing the duality diagram for ecologists. *J Stat Softw* **22**:1–20 (2007).

43 Peters NCB, Froud-Williams RJ, Orson JH The rise of barren brome Bromus-Sterilis in UKcereal crops. Brighton Crop Protection Conference: Weeds - 1993, Vols 1-3; 1993.

44 Benvenuti S, Weed seed movement and dispersal strategies in the agricultural environment. *Weed Biol Manag* **7**:141–157 (2007).

45 Colbach N and Sache I, Blackgrass (*Alopecurus myosuroides* Huds.) seed dispersal from a single plant and its consequences on weed infestation. *Ecol Model* **139**:201–219 (2001).

46 Naylor REL, Biological flora of the British Isles. No. 129 *Alopecurus myosuroides* Huds. (*A. agrestis* L.). *J Ecol* **60**:611–622 (1972).

47 Bradley R. Machinery hygiene to reduce grass weed spread. Available: https://www.fginsight.com/vip/vip/machinery-hygiene-to-reduce-grassweed-spread-228302017

48 Freckleton RP, Hicks HL, Comont D, Crook L, Hull R, Neve P *et al*., Measuring the effectiveness of management interventions at regional scales by integrating ecological monitoring and modelling. *Pest Manag Sci* **74**:2287–2295 (2018).

49 Naylor REL, *Alopecurus myosuroides* Huds. *J Ecol* **60**:611–622 (1972).

50 Brenchley WE, The weeds of arable land in relation to the soil on which they grow III. *Ann Bot* **27**:141–166 (1913).

51 Barnett DT, Stohlgren TJ, Jarnevich CS, Chong GW, Ericson JA, Davern TR *et al*., The art and science of weed mapping. *Environ Monit Assess* **132**:235–252 (2007).

52 Lambert JP, Childs DZ and Freckleton RP, Testing the ability of unmanned aerial systems and machine learning to map weeds at subfield scales: a test with the weed *Alopecurus myosuroides* (Huds). *Pest Manag Sci* **75**:2283–2294 (2019).

53 Lambert JPT, Hicks HL, Childs DZ and Freckleton RP, Evaluating the potential of unmanned aerial systems for mapping weeds at field scales: a case study with *Alopecurus myosuroides*. *Weed Res* **58**:35–45 (2018).

54 Morisette JT, Jarnevich CS, Ullah A, Cai W, Pedelty JA, Gentle JE *et al*., A tamarisk habitat suitability map for the continental US. *Front Ecol Environ* **4**:11–17 (2006).

55 Herrmann J, Hess M, Strek HJ, Richter O and Beffa R, Linkage of the current ALS-resistance status with field history information of multiple fields infested with blackgrass (Alopecurus myosuroides Huds.) in southern Germany. In Proc. 27th German Conference on Weed Biology and Weed Control (eds Henning, N. & Ulber, L.) Julius-Kühn-Archiv **452**: 42–49 (2016). https://doi.org/10.5073/jka.2016.452.006