Reciprocal Environmental Decision Support (REDS): better tailored advice in return for data

Ben Kenward[1,2], Nick Casey[1,3], Perline Bastid[1], Francis Buner[4], Ishi Buffam[5], Julie Ewald[1,4], Robert Kenward[1,6]

Institutions and addresses

[1]European Sustainable Use Group, Rue de la Science 10, Bruxelles 1000, Belgium

[2]Centre for Psychological Research, Oxford Brookes University, Headington Campus, Oxford, OX3 0BP, UK

[3]Anatrack Ltd, 52 Furzebrook Road, Wareham, Dorset, BH20 5AX, UK

[4]Game & Wildlife Conservation Trust, Burgate Manor, Fordingbridge, Hampshire, SP6 1EF, UK

[5]Department of Landscape Architecture, Planning and Management, Swedish University of Agricultural Sciences, Alnarp, Sweden

[6]UK Centre for Ecology and Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB, UK

Corresponding author

Ben Kenward
ben.kenward@wolfson.oxon.org

Data availability statement

Data is available via Supporting Information.

Acknowledgements

Thanks to Nicholas Aebischer and Johan Vegelius for useful discussions, to Prolific-recruited participants who went the extra mile, and to Bridget Kenward for much practical assistance. This study was funded by the EU Horizon grant PRO-COAST (Project ID 101082327).

Author contributions

Conceptualisation: BK, NC, JE, RK; Data curation: BK, NC; Analysis: BK; Funding acquisition: JE, RK; Methodology: all authors; Software: NC, BK; Writing (original draft): BK; Writing (review): all authors.

Conflict of interest statement

Ben Kenward, Nick Casey, and Robert Kenward are co-directors of Anatrack Ltd, a company that assisted in the development of technical systems used in the study, and Nick Casey is an employee. Following a decision at the 2025 board of directors meeting, the company is in the process of restructuring to a non-profit organisation.

Lay Summary

A summary for non-specialists is available at https://bit.ly/reds_brochure.

Reciprocal Environmental Decision Support (REDS): better tailored advice in return for data

Abstract

1. Environmental Decision Support systems use models to make tailored predictions about local circumstances based on user input. Citizen Science systems accept contributions about local circumstances to improve models. Each system type emphasises data flow in one direction: system to user or user to system, respectively.
2. We introduce a new system type that synergises data flows in both directions. Reciprocal Environmental Decision Support (REDS) systems process user-contributed information to (1) provide tailored predictions and (2) improve the model that generates the advice. We assessed existing decision support systems and found no REDS system for environmental management.
3. For our proof-of-concept REDS system, named Garden Advice, we began with an initial Bayesian species-habitat association model (for the House Sparrow *Passer domesticus*) and solicited data from UK residents about habitat structure and sparrow observations in domestic gardens. Based on habitat data, the model made predictions about sparrow presence, and effects of planned changes in garden management. Based on contributed habitat and sparrow data, it updated model parameter estimates, generally becoming more confident. One salient parameter update (a positive association of sparrows with grass) most likely reflects reality, and another (a reduction in the positive association with roof proximity) most likely reflects observation bias.
4. Cross-validated comparison indicated that the final model predicted the solicited data better than the original model. Thus, untrained observers can provide data of sufficient quality to refine a model already parameterised with trained observer data. Notwithstanding important questions about distinguishing observation bias from ecologically meaningful information, our system indicates that important synergies can be obtained from the REDS approach – later users of the system obtained better advice thanks to contributions from earlier users and automated Bayesian updating.

Keywords

Reciprocal environmental decision support; REDS; environmental decision support; EDS; Citizen Science; House Sparrow *Passer domesticus*; ecological modelling; Bayesian updating.

Word counts

281 (abstract); 7,967 (all other text)

Introduction

Introducing Reciprocal Environmental Decision Support (REDS)

A salient feature of many information systems is that query input data is not discarded once query output has been returned. Indeed, much of the modern information economy is built upon the value of query input (Bondia-Barceló et al., 2016). Although such use is often of ambiguous social benefit, raising concerns about privacy (Martin & Murphy, 2017) and marketing ethics (Frick et al., 2021), there are also examples representing unambiguous social good. For example, web search data reveals important trends in public consciousness of environmental issues (Nghiem et al., 2016), enables epidemics to be rapidly tracked (Ginsberg et al., 2009), and allows improvement of search algorithms (Cao et al., 2021).

Environmental decision support (EDS) systems typically require complex query input, because optimised predictions about environmental consequences of decisions require the situation to be described in some detail. In some cases, this will involve the input of public data, but in other cases, a land manager must collate and input detailed information about their land and its management that is not otherwise available to scientists or decision makers. The input data is processed to produce output that can support decisions but is not put to further use. As one example illustrating this general point, the Cool Farm Tool uses extensive farm activity data to estimate farm carbon footprint and to support decisions with analyses of hypothetical activity changes; it has been used for over 125,000 assessments (Cool Farm, 2025a; Kayatz et al., 2019). This implies a rich central database, but as they note, “the Cool Farm Alliance has not used any data stored in the tool” for purposes except returning assessment results (Cool Farm, 2025b).

The richness of EDS query input data, along with the lack of evidence for its use (beyond returning query output), implies that important opportunities are being missed. The typically multivariate nature of this data implies that it could parameterise predictive models that themselves could be useful within EDS. This would begin to bridge the gap between EDS and Citizen Science (CS). This work introduces a new type of system that combines the two (see conceptual groundwork in Kenward et al., 2013). Whereas CS emphasises the one-way data flow from citizen to system, and EDS emphasises the one-way data flow from system to citizen, Reciprocal Environmental Decision Support (REDS) represents a novel system type which closes this loop (Figure 1).

On a theoretical level, information loop closure has the potential to revolutionise the quality of data, models, and decisions, through iterative mutual improvement via feedback loops (Wiener, 1948). Here, we empirically demonstrate a REDS system that uses input not only to generate model predictions that can support decision making, but also to update the model itself. Thus, the system updates and increases confidence in model parameters as part of the same overall process that supports decisions based on model predictions. We illustrate this by implementing a habitat association model for an animal indicator-species offering predictions about species occurrence based on user-provided habitat descriptions; model updating is enabled by user-provided animal observations.

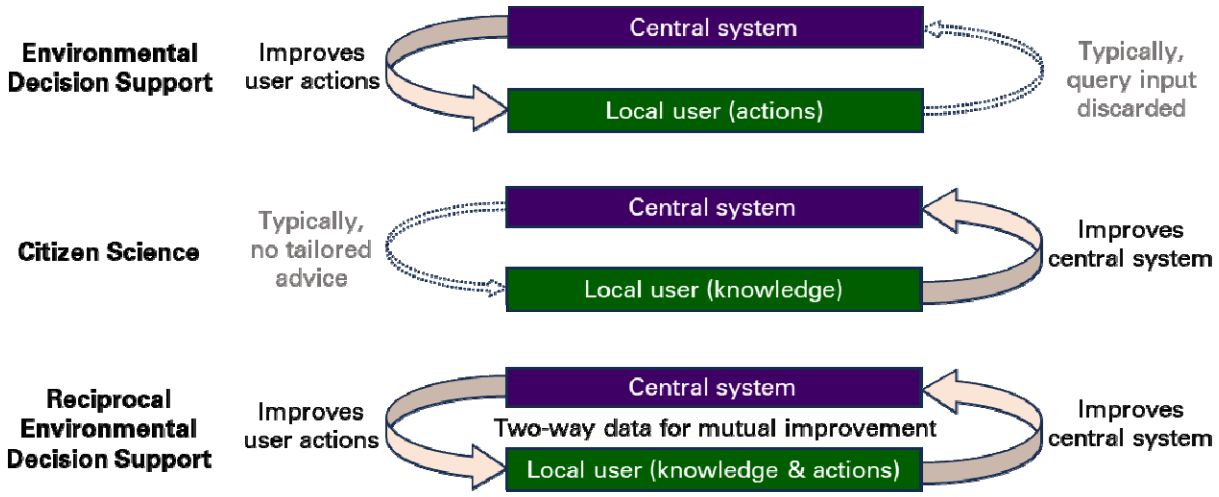


Figure 1. Information flows in EDS, CS, and REDS.

Environmental decision support

We define EDS systems as those with the primary purpose of supporting environmental management decisions and that provide model-generated predictions about environmentally relevant consequences of possible actions. Some systems such as databases can support environmental decisions without making predictions. In line with previous literature requiring EDS systems to feature “analysis” or “models” or to be “intelligent” (McIntosh et al., 2011, p. 1391), we exclude database-only systems from our definition. Predictions about unobserved conditions are key for EDS because they are the primary basis for evaluating the consequences of potential actions. REDS systems are EDS systems that are reciprocal; not only is user knowledge updated by the model predictions, but model predictions are updated by user observations (Figure 1).

Some systems include model predictions that might sometimes inform environmental decisions, although they are not primarily intended as EDS – an example is iNaturalist, primarily used for species monitoring (iNaturalist, 2025; Mason et al., 2025). Discussion of such systems is relevant because they may feature model updating that is reciprocal by our definition, and as such may be the most REDS-like available systems.

We define a REDS system as self-contained (though network-accessed) and with model updating on a time scale of a few days at most after data contribution. Otherwise, any EDS system informed by users in an indirect way (e.g., by a CS study reported in the literature) might be classified as REDS. Indirect and delayed closure of the information loop cannot lead to dynamic iterative updating, and is less likely to lead to productive synergies between CS and EDS.

Table 1 lists example systems and evaluates them as REDS candidates. Some systems feature reciprocal sharing of environmentally relevant data but are not REDS because there is no model to be updated, implying they are also not EDS. The PurpleAir system for sharing air quality measures is such an example (PurpleAir, 2025; Robinson et al., 2023). The species-mapping functionality of iNaturalist does have a model, and it is reciprocally updated: users report species observations that update a spatial model predicting occurrence at unobserved locations (Cole et al., 2023). The data flow and model updating is REDS-like, but iNaturalist is not in itself an EDS. The species occurrence models are based on spatial coordinates at very coarse scale, and

(as implemented) do not involve environmental conditions predictors, so the models are not suitable to support the what-if analyses needed for EDS.

Google Maps route-planning is presented as supporting environmental decisions, as it offers information on the relative energy efficiency of different routes. Further, it is reciprocal, as measurements of traffic speed from users update its models of traffic conditions which in turn inform route recommendations. Thus, Google Maps route-planning is the only existing REDS of which we are aware (although perhaps not a successful implementation, given issues with energy efficiency algorithms, Burris et al., 2023). It does stretch the definition, given it assists travel choices rather than ecosystem management directly.

Table 1: Data flows and REDS status for example candidate systems

	User → system	System → user				
	Data	Data	Model predictions	EDS	Reciprocal model update	Ecosystem management
PurpleAir	Air quality	Air quality	-	No	No	No
iNaturalist	Species occurrence	Species occurrence	Location → species occurrence	No	Yes	No
Google Maps	Traffic speed	Traffic speed	Route choice → energy efficiency	Yes	Yes	No
InVEST	Diverse variables	-	Diverse variables → ecosystem services	Yes	No	Yes
Garden Advice	Habitat, species occurrence	-	Habitat → species occurrence	Yes	Yes	Yes

Note: Ecosystem management denotes whether there are model predictions based on environmental condition variables, as opposed to (for example) purely spatial variables, which cannot be managed.

None of the above systems feature models where the inputs include environmental conditions (biotic or abiotic), although such models are key for most EDS systems, which focus on ecosystem management, and thus on manipulation of environmental conditions. Following a systematic review, we found no existing REDS systems with models based on environmental conditions – reciprocal model updating is common for species identification, but not for EDS (see Supporting Information). In general, EDS depends on models with environmental conditions predictors – for example, the well-known InVEST suite of models can provide predictions about a wide range of ecosystem services based on variables such as land cover (Natural Capital Project, 2025). However, we found no such EDS where user observations update a predictive model. We therefore introduce Garden Advice as the first REDS suitable for ecosystem management decisions.

The Garden Advice REDS system and our associated research questions

Garden Advice (GA) records (amongst other things) habitat descriptions and House Sparrow *Passer domesticus* observations in domestic gardens across the UK. We chose a habitat association model for sparrows for testing REDS-related research questions for four reasons. Firstly, house sparrows are in decline, including in the UK, for complex reasons, and they are frequently studied, and are thus a useful indicator species (Hanson et al., 2020). Secondly, despite their decline they are relatively common in UK domestic gardens, thanks to their association with humans (De Laet & Summers-Smith, 2007). Thirdly, UK domestic gardens are an excellent arena for citizen science, thanks to the high popularity of gardening in the UK (RHS, 2025; Sandhaus et al., 2019; Sharma et al., 2019). Fourthly, a raw dataset is available for fine-scale sparrow habitat association in UK domestic gardens, thanks to the Glasgow House

Sparrow Project (GHSP; Matthiopoulos et al., 2019), enabling an initial model for updating via the REDS approach.

Our primary research question was whether we could successfully implement a functioning REDS system. We define success as confirmation of the following pre-registered hypothesis: after updating the initial model with user contributed data, the model will better predict the user data. In principle, this represents a modest criterion, because if there is any information in the user data absent from the initial model, then with enough data, logically the criterion must be met. In practice, when soliciting data over the internet from untrained members of the public, across an appreciable geographic range, there are reasons why these conditions may not apply, rendering the question non-trivial. Our focus is on predicting user contributed data rather than the data that parameterised the initial model, because user data is likely more representative of house sparrows across the UK (the initial data came from one city). A secondary purpose of this report is to describe user advice-seeking behaviours within Garden Advice, as an example of investigate how those managing land garden use a REDS system.

Our methods include no independent verification of accuracy of citizen contributions. This raises concerns about citizen science data quality (Aceves-Bueno et al., 2017; Backstrom et al., 2024; Johnston et al., 2023; Kosmala et al., 2016), known to be questionable in similar contexts without additional verification (Williams et al., 2015). We return to this issue in the light of our results.

Methods

Declarations

Extensive additional detail concerning all methods is provided in Supporting Information. Large Language Models were not used to generate natural language but were sometimes used to check style and grammar. They were used to assist with writing of code. The procedure received ethical approval from the [ANONYMISED] human-participant research-ethics internal assessment procedure. Participants gave fully informed consent in a screening questionnaire. We pre-registered a specific method to test the pre-registered hypothesis, which we followed except where unforeseen circumstances intervened, as noted where relevant.

Participants

Prolific, a service for social scientists to recruit paid participants, known for good quality data in comparison to other similar services (Albert & Smilek, 2023; Peer et al., 2022), was used in May 2025 to commission 71 UK residents to describe their gardens using Garden Advice. An additional 68 UK residents made initial attempts to participate, but could not, for example due to being screened out for misidentifying birds in the screening questionnaire that followed Prolific recruitment. All garden data obtained during the study period that logically and ethically could be included in model updating was included. A small number of broad requirements for participation (e.g., having an upper secondary education) were implemented in Prolific to optimise participant ability and motivation to succeed at mapping a domestic garden. Most participants were middle-aged; most were in full-time work; most had university degrees.

Garden Advice app

The system provides user-friendly methods for citizens to map habitats in domestic gardens and to specify areas where they typically observe House Sparrows (and other species, outside current

scope). The system is REDS because it also (1) provides users with scores indicating the likelihoods of observing sparrows in gardens they describe, and (2) updates the model that generates these scores with the data that users provide. The system also provides model scores indicating how changes to the garden would impact the likelihood of observing sparrows. Thus, the system is designed to be capable of providing better EDS, the more it is used – the hallmark of the REDS approach. Garden Advice can be accessed at https://garden.earthadvice.org (those providing non-genuine data for testing purposes are asked to enable testing mode).

The user interface for Garden Advice is characterised by two primary functionalities: (1) drawing polygons on a base map (Google Maps satellite view); (2) obtaining scores representing predictions from models that process the entered map data. Users can enter polygons representing habitats or areas where different animal species are typically observed. Additional functionality includes the ability to enter planning mode. Planning mode copies the existing map to a new version designed to allow what-if edits permitting model score comparisons between the real map and the modified plan.

In-app instructions guided the current users to (1) find their garden on the base map, (2) draw a boundary around it, (3) add habitat patches until the bounded area was filled, (4) certify that the map was accurate (with direction to editing functions if not), (5) answer whether they had ever seen a House Sparrow in their garden, (6) if they answered yes, draw polygons around all specific areas where sparrows have been seen. All data was contributed without any non-automated assistance to users.

Garden Advice also incorporates a carbon model estimating how much carbon gardens sequester. It provides these estimates in the form of hours of human life saved, by converting carbon to human mortality via the thousand-ton rule (see Pearce & Parncutt, 2023 for a presentation of this rule, and Bressler, 2021 for an example model from which it is derived). Although not part of the REDS functionality, this is an important part of Garden Advice because it can attract users (this was how we developed the system), and because framing carbon sources and sinks in terms of countable human deaths and lives saved is a useful messaging tool (Kenward, 2024).

The sparrow habitat association model

Variables

Model variables were chosen for compatibility with the publicly available Glasgow House Sparrow Project dataset (GHSP), consisting of observations from trained surveyors at 32 individual colony areas in and around the city of Glasgow (Matthiopoulos et al., 2019). Each data point represents a 2 m^2 map grid-cell where a sparrow was (1) or was not (0) observed. Predictor variables are proportion of cell covered by each of six habitat types (tree, bush, hedge, grass, roof, and artificial surface) plus two distance variables representing the distance from the cell to the nearest hedge or roof. Garden Advice (GA) allows mapping of two additional habitat types anticipated to be necessary for participants to fully map their gardens: water and cultivated beds. In the current model, cultivated beds are treated as bushes and water is ignored.

The polygon vector format of GA internal data is converted to the grid cell format to allow modelling of the GA and GHSP data combined. The binary dependent variable is 1 if any sparrow observation area overlaps with the cell, otherwise 0. Utilised GHSP distances were approximate reconstructions because available GHSP distances were z-transformed before release; GA roof distances are calculated using pre-existing map data as well as user data; GA

hedge distances are imputed as the maximum observed value in gardens with no hedges (96.7 m). For presentation of analysis, distance variables are converted to proximity variables by reversing the sign, so that effects have a more intuitive direction.

Model specification and updating

The model is a Bayesian binary logistic regression, with R syntax:

```
sparrow ~ hedge + grass + tree + roof + artificial + bush + hedgeProximity + roofProximity +
    (1 | location) +
    (0 + hedge + grass + tree + roof + artificial + bush + hedgeProximity + roofProximity || location)
```

In other words, there are eight fixed predictors, an overall intercept, a random intercept for each observed location (GHSP colony or GA garden), and random slopes for each fixed predictor and location (uncorrelated to ease model fit). The random slopes control for intra-location spatial autocorrelation, as well as capturing any inter-location heterogeneity in habitat associations.

The model was initially fit to all GHSP data and incorporated into the GA system. It was then updated with the new GA data every day during the study period, entirely automatically except for manual addition of certain data contributions to an exclusion list (Table S1).

GHSP data was sampled in case-control format – each colony area was observed until a sparrow was registered at 40 grid cells, then 40 grid cells where no sparrow was observed were randomly selected, and habitat was described for the 80 cells. GA data, on the other hand, comprises full presence-absence information, enabling prevalence calculations. We used a prevalence-based offset to make the data sets compatible.

Pre-registration stated that models would be fitted using the standard Markov chain Monte Carlo (MCMC) approach with the R package rstanarm (Goodrich et al., 2025). However, we had underestimated how run time for MCMC Bayesian updating would scale up with additional data. It became evident that leave-one-group-out cross-validation (requiring one model to be constructed per garden or colony; see below) was only practical with a faster alternative fitting method – we used Integrated nested Laplace approximation (INLA) via the R-INLA package (Rue et al., 2017). Other analyses are presented using both approaches. Priors were the default in all cases (as pre-registered) except for random slope variance hyperparameters in the INLA models, where initial failures to fit, combined with retrospective consideration of appropriate prior beliefs, led to a specification of a comparatively high variance prior.

Model predictions and their evaluation

Preliminary investigations indicated that random components of any version of the model explained large amounts of variation. However, this variation cannot be used for predictions at novel locations, and we therefore only use fixed model components for making predictions.

For evaluating prediction accuracy, we pre-registered the use of Brier scores and the relevant decompositions. Brier scores are equivalent to mean squared error (MSE; this name is common in forecasting contexts). Brier scores are associated with standard methods for decomposition into different components (Siegert, 2017). The score and its decompositions all range between 0 and 1. The signs and names can be counterintuitive. We retain the standard sign directions, but adopt later clearer names given to the decompositions (Merkle & Hartman, 2018). Discrimination (i.e., resolution) measures how well separated the forecast probabilities are for observed presences versus absences (higher scores are better). Miscalibration (or calibration)

measures how far the forecast probabilities deviate from the observed frequencies on average (lower scores are better). We ignore the third and final decomposition, uncertainty, which relates only to uncertainty in the data and so does not assist with model comparison.

Our primary model comparison examines the relative abilities of the initial GHSP-data-only model, and the model updated with GA data, to predict GA data. This is because the GA data is assumed closer to the population we aim to represent with this model – but we also explore other predictions and discuss this assumption. When comparing model performance metrics, we generate the metrics using leave-one-group-out analysis (Vehtari, 2025), a hierarchical leave-one-out cross-validation where a model is never evaluated for predicting data it is trained with.

When comparing the initial GHSP-only model with the final GA-updated model, we deviate from pre-registration by recreating the GHSP-only model with an offset calculated from the final GA prevalence data. The effectively gives the GHSP-only model an unrealistic advantage, because the GHSP-only model then incorporates the prevalence data present only in the GA data. We chose this because otherwise any model calibration improvement following updating might reflect nothing more than the trivial fact that the GHSP-only model is uncalibrated for prevalence due to the GHSP case-control data sampling method.

Results

Data description

The median size of holdings mapped in Garden Advice (GA) was 509 m^2, first to third quartiles 199 m^2 to 1257 m^2, mean 1299 m^2 (including mapped buildings). Gardens were distributed throughout the United Kingdom (Figure 2). Descriptive statistics for model variables are presented in Table 2. This includes GA sparrow prevalence data, but prevalence cannot be reconstructed for GHSP because of case-control sampling. Examples of user-created maps are shown in Figure 3.

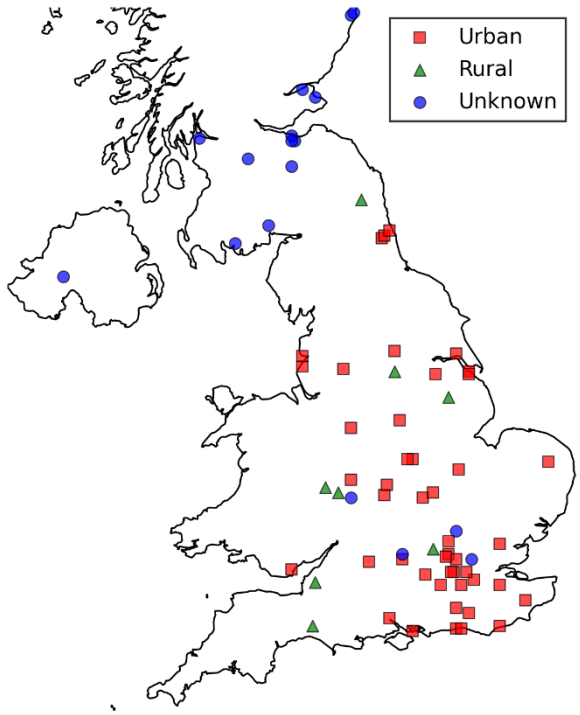


Figure 2. Locations of the 71 Garden Advice (GA) gardens within the UK, snapped to a 5 km grid for privacy. Rural / urban classification data (ONS Geography, 2025) covers England and Wales only and is sometimes unavailable for other reasons.

Table 2. Descriptive statistics for model variables. GA and GHSP are most appropriately compared using cells with no sparrows, given the different sampling approaches.

	All map cells			Map cells with no sparrows					
	GA			GA			GHSP		
Variable	Mean	Median	Present	Mean	Median	Present	Mean	Median	Present
Cell percentage cover									
Sparrow	22%	17%	85%	-	-	-	-	-	-
Roof	12%	4%	56%	13%	4%	56%	22%	20%	100%
Grass	28%	26%	83%	25%	24%	80%	25%	23%	97%
Bush	17%	14%	82%	15%	9%	77%	7%	4%	78%
Tree	16%	7%	61%	15%	6%	58%	5%	2%	50%
Hedge	3%	0%	20%	3%	0%	17%	7%	6%	78%
Artificial	23%	20%	85%	26%	20%	85%	33%	33%	100%
Distance to nearest (m)									
Roof	8.9	6.2	100%	8.5	5.7	100%	11.4	11.1	100%
Hedge	14.2	16.0	20%	15.6	16.8	20%	13.4	12.8	100%

Notes: For cover variables (including sparrow observation areas), present means the percentage of locations (gardens or colonies) with non-zero cover. For distance variables, present means the percentage of locations where the distance variable was available.

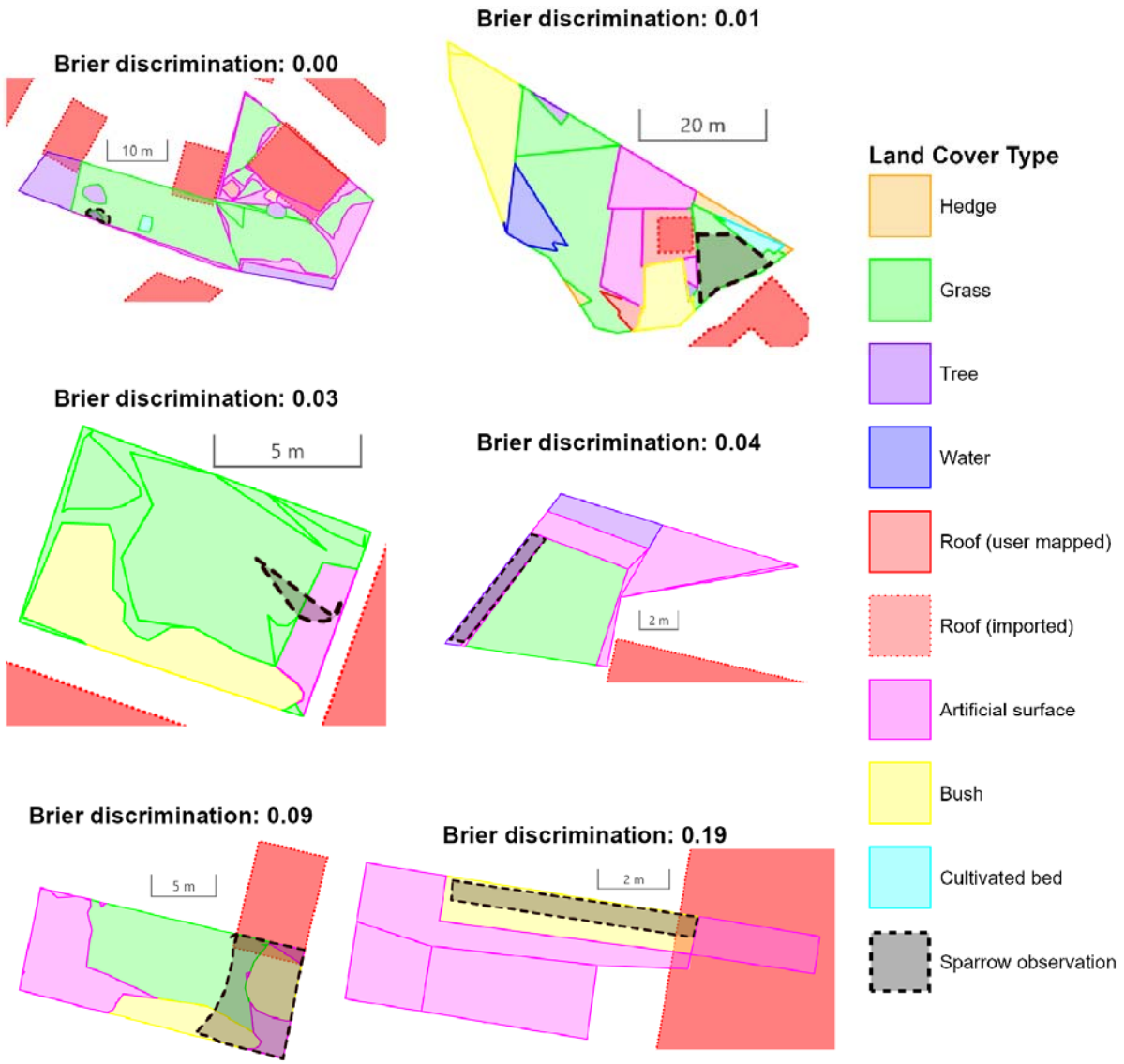


Figure 3. Six user-provided GA garden maps, selected as those nearest the duodeciles (0%, 20%, … 100%) of the Brier score discrimination decomposition distribution, to provide examples of where sparrows are poorly and well predicted by the final model. Dotted edges indicate buildings not drawn by users but automatically imported to calculate roof proximity.

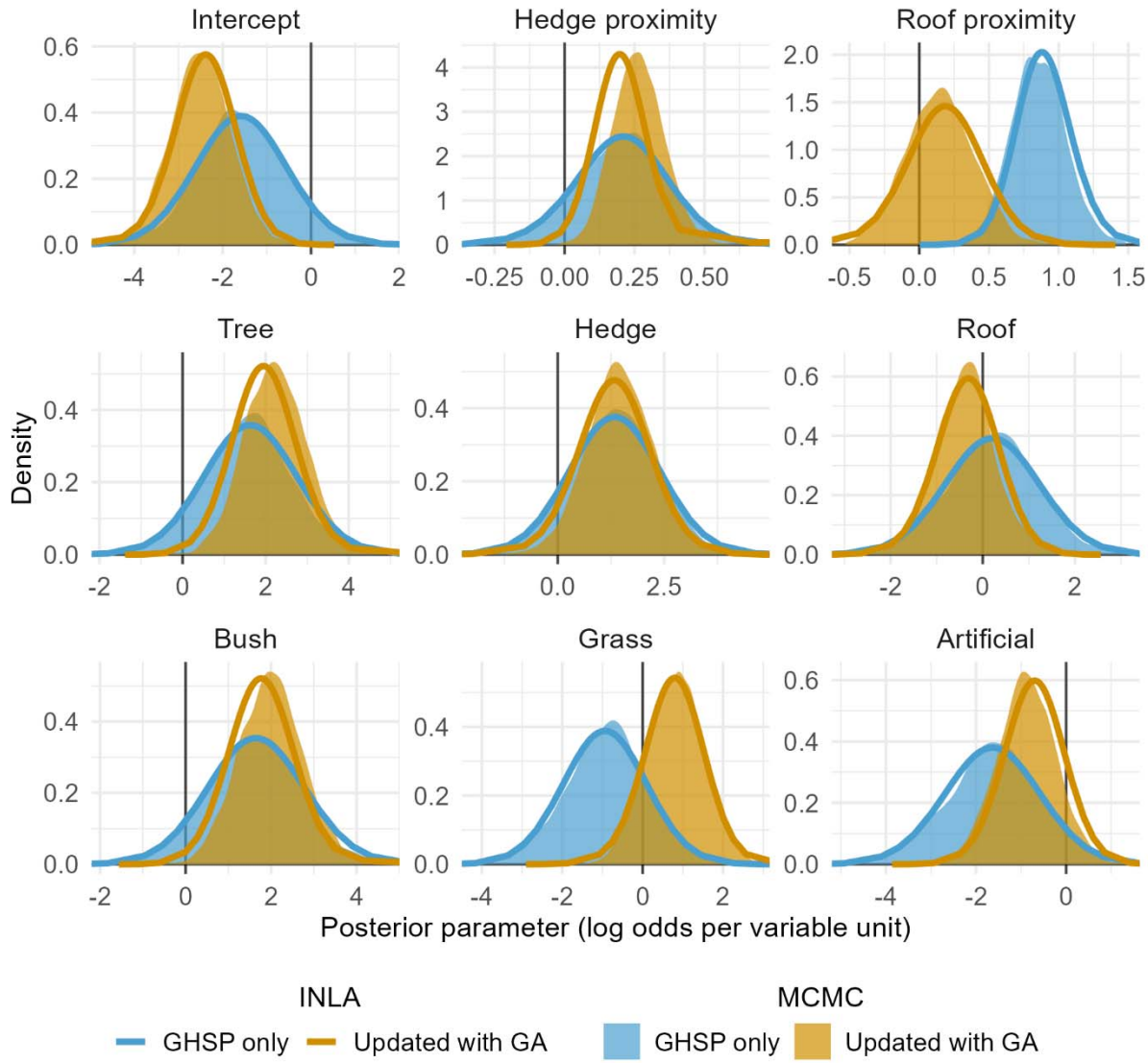
Effect of model updating on model parameters

Figure 4 shows how fixed model parameters were updated. For trees, bushes, and hedges (cover and proximity), the model became more confident of more positive effects. The effect size differences were small, but the model updated from being only weakly confident to being very confident in positive effects. There was a salient change for grass, where the model updated from

being weakly confident in a negative effect to slightly more confident in a positive effect (with a 91% probability this represents a real difference[1]).

For roofs and artificial surfaces, on the other hand, the model decreased its estimates of impact (positive and negative, respectively). Here, the most salient change was for roof proximity, with a 98% probability that this represents a real difference. This change also implied a reduction in confidence about the effect (unique to this parameter).

Figure 4 also indicates that INLA approximations converged very well with MCMC models, with one exception – the hedge proximity parameter. As even this parameter converged reasonably well, INLA appears a generally very good approximation in this context. The imperfect convergence for hedge proximity likely relates to greater difficulty fitting hedge parameters due to the absence of hedges from most GA gardens.



[1] This probability is derived as the probability that a point from the first posterior differs from a point in the second posterior in the same direction as the observed difference in medians.

Figure 4. Posterior distributions for model parameters before and after updating of GHSP-data model with GA data. Variable units are metres or cell proportion covered, as appropriate.

Effect of model updating on predictive power

Updating with GA data improved the model's overall ability to predict the GA data, but decreased its ability to predict the GHSP data, and thus ability to predict all data together was unchanged (Brier scores in Figure 5).

The Brier score can be decomposed into two separate model accuracy properties (Merkle & Hartman, 2018; Siegert, 2017). Discrimination reflects the ability to correctly rank observed cells by their probability of containing a sparrow observation, and (mis)calibration reflects the degree to which predicted probabilities match observed frequencies, i.e., whether predicted and observed occurrence rates match. The trends for both decompositions are in the same direction as the overall Brier score: improvements for GA data prediction, worsening for GHSP data prediction, and no change overall. Due to greater uncertainty when decomposing (credible intervals in Figure 5), it's not possible to determine whether the overall improvement relates primarily to discrimination or miscalibration or both. The only relevant pre-registered hypothesis, which predicted improvements in overall Brier score and its decompositions when predicting GA data, was thus confirmed in essence.[2]

[2] There was one additional pre-registered hypothesis, that the updated model would explain variance; this hypothesis is ignored as it is trivial given that the primary hypothesis is confirmed.

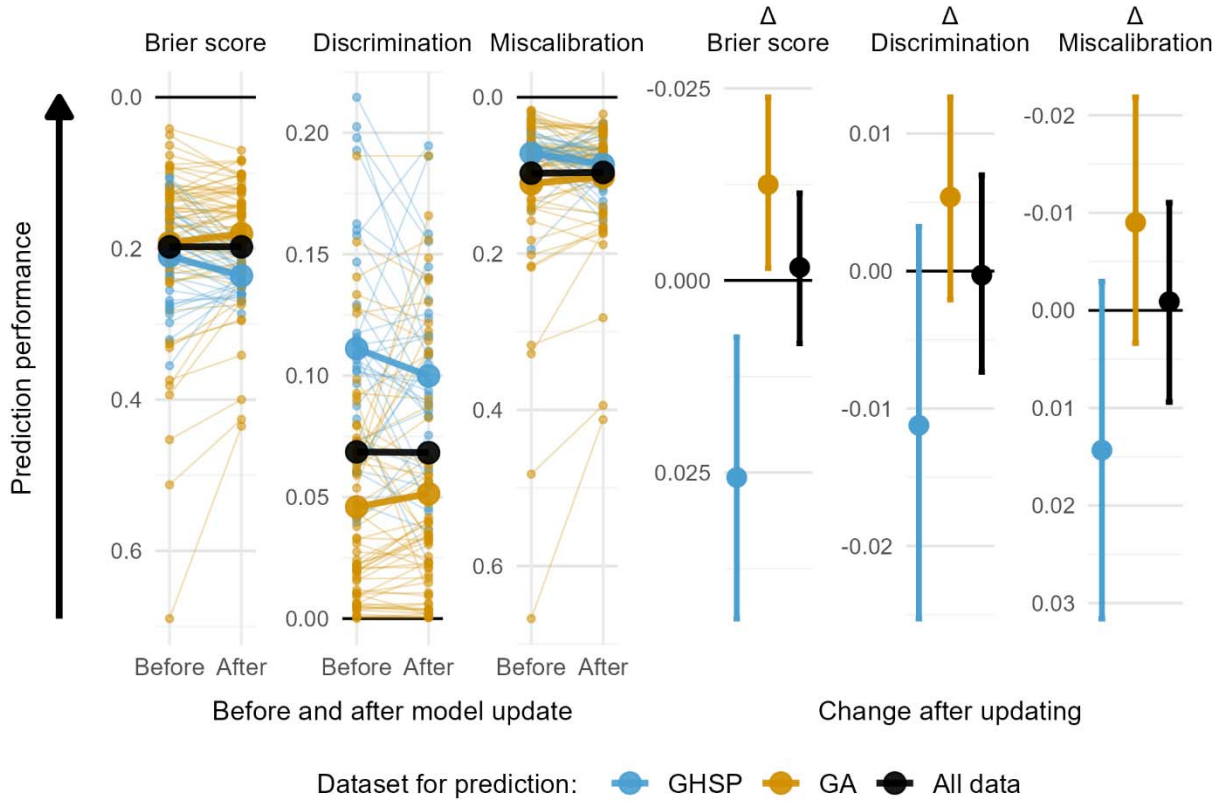


Figure 5. Model performance before and after the GHSP-only model was updated with GA data, when predicting each data set, and both. Vertical scale direction is set so upwards always represents improvement. The three left-hand panels show overall Brier score and decompositions. Small points and lines show performance for individual gardens and colonies, while large points and lines show mean performance across each prediction dataset. The three right-hand panels show median and 95% credible intervals for the change in each metric after model updating. Leave-one-out cross-validation is applied so that prediction evaluation is always for locations not in training data.

For the current EDS application, thresholding model output to produce a presence-absence prediction is unnecessary, but in many contexts, decision making requires a best guess. We therefore explored presence-absence predictions for GA data with thresholds chosen according to various criteria, with the comparisons finding at least small improvements for the updated model, depending on threshold selection approach.

Effect consistency

Finally, we consider an aspect of model prediction particularly relevant in decision-making contexts – effect consistency, defined as the estimated probability that any random garden will display an effect in the predicted direction. This is correlated with but separable from the estimated fixed effect, because it also depends on effect heterogeneity – the extent to which the same effect tends to apply across random locations. Figure 6 illustrates good confidence in most vegetation-related effects and effect consistencies – for example, the positive effect of trees is predicted to apply in approximately 80% of gardens in which they occur. There is some weak

evidence that the consistency of hedge effects might be slightly greater than expected, given their fixed effect sizes, although trees and bushes are estimated to be the most beneficial vegetation. The equivalent figure for the GHSP-only model (Figure S2) indicates that GHSP data by itself does not enable confident predictions except for those based on roof proximity and artificial surface.

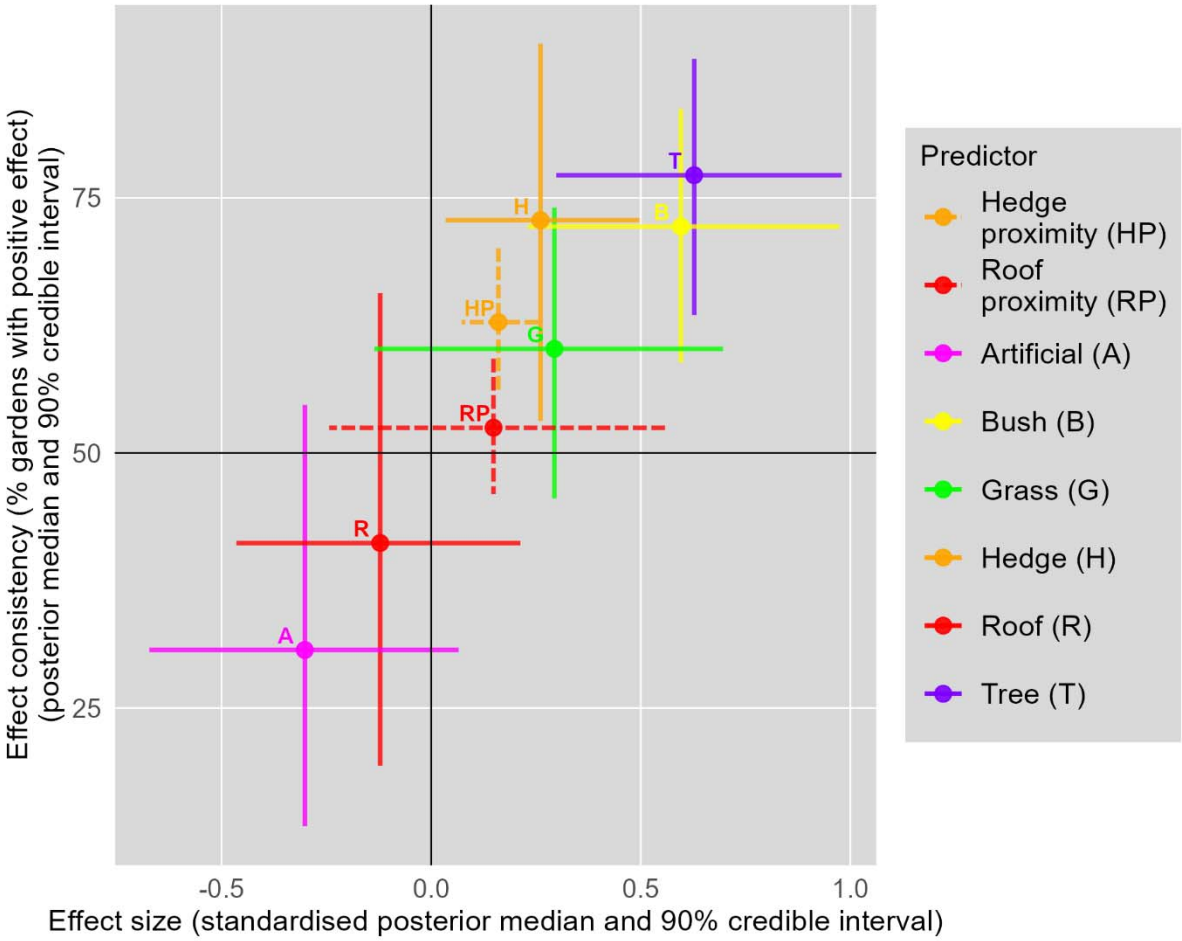


Figure 6. Effect consistency and standardised effect size across predictors for GA-updated MCMC model. Effect sizes are standardized by multiplying regression parameters by the pooled within-location standard deviation for each predictor.

Decision-support-seeking behaviours

Participants were paid small sums to follow specific instructions within the GA system, necessary to provide data that could support model updating. They were not instructed to obtain environmental decision support by running models to see their garden's scores for sparrows or carbon. However, 56% examined their sparrow scores, 59% examined their carbon scores, 76% examined at least one of those, 18% planned changes to their garden using planning mode, and 11% examined how these changes affected scores.

Discussion

Observed model improvements

The REDS approach was successfully demonstrated in a proof-of-concept system, Garden Advice (GA). Users contributed local habitat and species occurrence data, and this refined the decision support available to them, as it caused the model generating predictions to become more accurate. The overall Brier score improvement of a little over 0.01 corresponds to probability predictions improving by a little over 0.1 on average – a non-trivial improvement – and threshold-based approaches also revealed meaningful improvements. As discussed in Supporting Information, the Brier score improvement almost certainly represents an improvement in discrimination, because our model offsetting approach makes calibration improvement unlikely. Whether these model improvements reflect reality depends on the veracity of the contributed data, but prior knowledge suggests that at least one change to the model – an increase in association with grass – represents real ecological information.

This was the most marked parameter change at model updating. A weak confidence in a negative association for grass changed to weak confidence in a positive association. This update is consistent with other evidence on House Sparrow habitat utilisation. Anecdotal (Bradbury, 2024; Teo, 2023) and systematic (Vincent, 2005) observations of House Sparrows suggest that although grass is not a preferred foraging habitat, they can obtain seeds and invertebrates there. In Spain, reduction of grass in urban parks has been associated with House Sparrow decline (Bernat-Ponce et al., 2020). However, other studies also indicate negative associations with grass (Havlíček, 2021; Weir, 2015). One likely resolution to this conflicting evidence relates to temporal variation. House Sparrow preference for grass has been observed to fluctuate month-on-month and year-on-year (Vincent, 2005), so grass may be preferred when temporary conditions increase its relative benefits. All data here was from May 2025, but REDS systems are well suited to provide data over time and across circumstances to support a richer model that could address issues of seasonality and other reasons for heterogenous effects.

Of the remaining changes in model parameters, most represented increased confidence in slightly larger positive effects. This was the case for trees, bushes, and hedges (cover and proximity), and this is also supported by the general literature. There were weak indications that the effects of hedges may be more consistent across locations than would be predicted by effect size alone. This accords with previous observations of the general importance of hedgerows for birds and reinforces the importance of conserving hedges (Glasgow House Sparrow Project, n.d.; Hinsley & Bellamy, 2000). Such effect consistency estimates are only available when sampling large numbers of independent locations, as they are transformations of inter-location effect heterogeneity. REDS and other citizen science systems are well suited to collecting such data (Garretson et al., 2023). Effect consistency is important for EDS as it is appropriate for decision support to emphasise consistent effects, independently of effect size, because such effects are robust to local idiosyncrasies. We are unaware of previous uses of our specific method for intuitively presenting effect heterogeneity, by transforming to the percentage of locations where the effect is expected to apply. We suggest this presentation method could be generally beneficial for decision support.

The GHSP-only model's initial strong confidence in the association of sparrows with roof proximity was removed by the GA data model updating. This is unlikely to be realistic, given the close affiliation of House Sparrows with human dwellings due to commensality and nesting

(Shaw et al., 2008). This discrepancy highlights a key issue: the accuracy of Garden Advice user observations was not determined, and this result likely reflects observation bias. Birds on or near roofs are harder to see, especially for those observing through windows from indoors. Another study that also solicited non-expert animal observations from UK domestic gardens found that although participants could provide accurate habitat descriptions, hedgehog presence-absence reports were uncorrelated with objective data (Williams et al., 2015). Hedgehogs are harder than sparrows for human observers to detect, but this nevertheless provides a sobering reminder of the necessity for independent confirmation of citizen science data.

REDS – broader implications

Reviews indicate that citizen science data is often reliable, and measures such as additional training provide benefits (Aceves-Bueno et al., 2017; Backstrom et al., 2024; Johnston et al., 2023; Kosmala et al., 2016). Further, although our study involved data solicitation from a wide public, in practice EDS systems are often used by land managers such as farmers, who might be better able to provide accurate observations. Nevertheless, the next stage in REDS concept validation must involve independent verification of data quality. This is not only a requirement because of the citizen science component of REDS – it is also relevant since model validation is one of the factors most associated with EDS system uptake (Walling & Vaneeckhaute, 2020).

Model predictions of species occurrence under local conditions may not be useful for users who can make their own occurrence observations. However, REDS systems tend to get more accurate for everyone, even those who don’t contribute observations necessary for model updating, providing that some users do. Further, for users who have made observations, predictions might prompt fruitful investigation of why observations are not as predicted. They can also be used for what-if analysis examining planned habitat changes – a feature inherent to most EDS and available within GA.

The REDS design concept joins that of the digital twin as a system using modern technology to close an information loop to improve model predictions (De Koning, 2025). REDS systems could also be digital twins, although REDS does not necessarily involve real-time modelling of an ongoing phenomenon – the defining feature of digital twins – and digital twins don’t necessarily directly support environmental decision making. Despite recent reviews of digital twins in environmental science (Hazeleger et al., 2024; Trantas et al., 2023), we are unaware of any digital twin system yet operating to offer practical decision support for ecosystem managers. These reviews list appreciable practical challenges for the digital twin approach – perhaps in some contexts, the less ambitious REDS approach could offer some of the same benefits more straightforwardly.

Future REDS systems would benefit from more complex models. Habitat association models do not consider causal mechanisms and do not necessarily predict long-term species health, because local observations may not predict longer term demographic changes (Bacon et al., 2017; Matthiopoulos et al., 2019; Regos et al., 2021). Experimental provision of food was found to have a large impact on numbers of fledged House Sparrow chicks, but a smaller impact on the adult population, indicating that foraging resources may not be the primary limiter of abundance (Peach et al., 2015). That said, reviews indicate that although there are important exceptions, the availability of preferred habitats does tend to predict species abundance (Boyce et al., 2016; Weber et al., 2017). In principle, the REDS citizen science approach could also track demographic change, as one route to support more complex and accurate models.

Further development of the REDS concept will benefit from closer examination of the potential of the social synergies generated by combining citizen science and EDS. There is always potential for useful synergy when closing an information loop, allowing feedback mechanisms to operate (Wiener, 1948). For REDS, some of these synergies might relate to human motivations to engage in reciprocal behaviour (Bowles & Gintis, 2011). It is likely that public data sharing that already exists on platforms such as iNaturalist and PurpleAir is motivated in large part by reciprocity (Metzger & Runge, 2023; Pai & Tsai, 2016). Our users were primarily motivated to engage with the system for a small payment, contingent on completion of observation reports, and they were never instructed to use the EDS functions. It's therefore notable that most of them did request model predictions, so later users did benefit from improved predictions due to data contributed by earlier users.

Examples of existing initiatives that could become REDS include membership-organisation citizen-science initiatives, such as the British Trust for Ornithology's Garden Birdwatch (Morrison et al., 2014), schemes associated with the Bumblebee Conservation Trust (Sharma et al., 2019), or the Game & Wildlife Conservation Trust's Partridge Count Scheme (Ewald et al., 2009). Participation could be increased, and relevant environmental management behaviour could be improved, as citizens receive tailored advice as an integral part of reporting their contributions. A concrete example from this study is that, thanks to the provision of participant data that updated the model to separate grass from artificial surfaces, participants could receive advice that lawns are better for sparrows than decking or paving.

Additional motivation for engagement may come from participants more directly seeing how individual contributions improve centralised knowledge. Stakeholder engagement is a key determinant of EDS system success (Walling & Vaneeckhaute, 2020). Provided that careful consideration is given to the structure of the initial model, REDS can save time for the scientists operating such schemes, as analysing and reporting the data is an automated and integrated part of the data-collection process itself. Like digital twins, the REDS concept is at an early stage of development, and the next stages must also focus on how this can be achieved without sacrificing the quality of analysis and conclusions.

In summary, the sense that one is helping at the same time as being helped could be a strong factor for further activating the synergies between EDS and citizen science that this study indicates REDS can provide. These synergies offer positive feedback loops potentially generating a virtuous circle of greater user engagement, data provision, advice quality, and advice uptake.

References

Aceves-Bueno, E., Adeleye, A. S., Feraud, M., Huang, Y., Tao, M., Yang, Y., & Anderson, S. E. (2017). The Accuracy of Citizen Science Data: A Quantitative Review. *The Bulletin of the Ecological Society of America*, *98*(4), 278-290.

Albert, D. A., & Smilek, D. (2023). Comparing attentional disengagement between Prolific and MTurk samples. *Scientific Reports*, *13*(1), 20574.

Backstrom, L. J., Callaghan, C. T., Leseberg, N. P., Sanderson, C., Fuller, R. A., & Watson, J. E. M. (2024). Assessing adequacy of citizen science datasets for biodiversity monitoring. *Ecology and Evolution*, *14*(2), e10857.

Bacon, L., Hingrat, Y., Jiguet, F., Monnet, A.-C., Sarrazin, F., & Robert, A. (2017). Habitat suitability and demography, a time-dependent relationship. *Ecology and Evolution*, *7*(7), 2214-2222.

Bernat-Ponce, E., Gil-Delgado, J. A., & López-Iborra, G. M. (2020). Replacement of semi-natural cover with artificial substrates in urban parks causes a decline of house sparrows Passer domesticus in Mediterranean towns. *Urban Ecosystems*, *23*(3), 471-481.

Bondia-Barceló, J., Castellà-Roca, J., & Viejo, A. (2016). Building Privacy-Preserving Search Engine Query Logs for Data Monetization. 2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld),

Bowles, S., & Gintis, H. (2011). *A cooperative species: Human reciprocity and its evolution*. Princeton University Press.

Boyce, M. S., Johnson, C. J., Merrill, E. H., Nielsen, S. E., Solberg, E. J., & van Moorter, B. (2016). REVIEW: Can habitat selection predict abundance? *Journal of Animal Ecology*, *85*(1), 11-20.

Bradbury, K. (2024). *The house sparrow (Passer domesticus)*. https://www.encounter-nature.com/features/the-house-sparrow-passer-domesticus

Bressler, R. D. (2021). The mortality cost of carbon. *Nature Communications*, *12*(1), 4467.

Burris, M., Khan, M., & Johnson, J. (2023). *"Is That Route Really the Most Fuel-Efficient?"* (National Institute For Congestion Reduction Research Reports, Issue 17). https://doi.org/10.5038/CUTR-NICR-Y3-4-8

Cao, K., Chen, C., Baltes, S., Treude, C., & Chen, X. (2021). Automated query reformulation for efficient search based on query logs from stack overflow. 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE),

Cole, E., Van Horn, G., Lange, C., Shepard, A., Leary, P., Perona, P., Loarie, S., & Mac Aodha, O. (2023). *Spatial Implicit Neural Representations for Global-Scale Species Mapping* Proceedings of the 40th International Conference on Machine Learning, Proceedings of Machine Learning Research. https://proceedings.mlr.press/v202/cole23a.html

Cool Farm. (2025a). *Cool Farm*. Retrieved 2025-10-03 from https://coolfarm.org/

Cool Farm. (2025b). *Cool Farm: Frequently Asked Questions*. Retrieved 2025-10-03 from https://coolfarm.org/frequently-asked-questions/

De Koning, K. (2025). The crane radar: Development and deployment of an operational eco-digital twin. *Ecological Informatics*, *85*, 102938.

De Laet, J., & Summers-Smith, J. D. (2007). The status of the urban house sparrow Passer domesticus in north-western Europe: a review. *Journal of Ornithology*, *148*(2), 275-278.

Ewald, J. A., Kingdon, N. G., & Santin-Janin, H. (2009). The GWCT Partridge Count Scheme: a volunteer-based monitoring and conservation promotion scheme. In S. B. Cederbaum, B. C. Faircloth, T. M. Terhune, J. J. Thompson, & J. P. Carroll (Eds.), *Gamebird 2006: Quail VI and Perdix XII* (pp. 27-37). Warnell School of Forestry and Natural Resources.

Frick, V., Matthies, E., Thøgersen, J., & Santarius, T. (2021). Do online environments promote sufficiency or overconsumption? Online advertisement and social media effects on clothing, digital devices, and air travel consumption. *Journal of Consumer Behaviour*, *20*(2), 288-308.

Garretson, A., Cuddy, T., Duffy, A. G., iNaturalist Citizen, S., & Forkner, R. E. (2023). Citizen science data reveal regional heterogeneity in phenological response to climate in the large milkweed bug, Oncopeltus fasciatus. *Ecology and Evolution*, *13*(7), e10213.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, *457*(7232), 1012-1014.

Glasgow House Sparrow Project. (n.d.). *Help your sparrows*. Retrieved 2025-12-15 from https://www.housesparrowscience.com/help-your-sparrows/

Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2025). *rstanarm: Bayesian applied regression modeling via Stan. R package version 2.32.1*. https://mc-stan.org/rstanarm/

Hanson, H. E., Mathews, N. S., Hauber, M. E., & Martin, L. B. (2020). The house sparrow in the service of basic and applied biology. *Elife*, *9*.

Havlíček, J. (2021). *The breeding and foraging ecology of the House Sparrow in rural and urban environments* [PhD Thesis, University of South Bohemia in České Budějovice].

Hazeleger, W., Aerts, J. P. M., Bauer, P., Bierkens, M. F. P., Camps-Valls, G., Dekker, M. M., Doblas-Reyes, F. J., Eyring, V., Finkenauer, C., Grundner, A., Hachinger, S., Hall, D. M., Hartmann, T., Iglesias-Suarez, F., Janssens, M., Jones, E. R., Kölling, T., Lees, M., Lhermitte, S., . . . Vossepoel, F. C. (2024). Digital twins of the Earth with and for humans. *Communications Earth & Environment*, *5*(1), 463.

Hinsley, S. A., & Bellamy, P. E. (2000). The influence of hedge structure, management and landscape context on the value of hedgerows to birds: A review. *Journal of Environmental Management*, *60*(1), 33-49.

iNaturalist. (2023). *Introducing the iNaturalist Geomodel*. https://www.inaturalist.org/posts/84677-introducing-the-inaturalist-geomodel

iNaturalist. (2025). *iNaturalist*. Retrieved 2025-10-03 from https://www.inaturalist.org/

Johnston, A., Matechou, E., & Dennis, E. B. (2023). Outstanding challenges and future directions for biodiversity monitoring using citizen science data. *Methods in Ecology and Evolution*, *14*(1), 103-116.

Kayatz, B., Baroni, G., Hillier, J., Lüdtke, S., Heathcote, R., Malin, D., van Tonder, C., Kuster, B., Freese, D., Hüttl, R., & Wattenbach, M. (2019). Cool Farm Tool Water: A global on-line tool to assess water use in crop production. *Journal of Cleaner Production*, *207*, 1163-1179.

Kenward, B. (2024). *The 1000-ton rule, which emphasises the mortality cost of carbon, can be an effective persuasion tool, including for use by activists.* https://www.benkenward.com/articles/the_1000_ton_rule_can_be_an_effective_persuasion_tool.pdf

Kenward, R., Papathanasiou, J., Manos, B., & Arampatzis, E. (Eds.). (2013). *Transactional Environmental Support System Design: Global Solutions*. IGI Global.

Kosmala, M., Wiggins, A., Swanson, A., & Simmons, B. (2016). Assessing data quality in citizen science. *Frontiers in Ecology and the Environment*, *14*(10), 551-560.

Martin, K. D., & Murphy, P. E. (2017). The role of data privacy in marketing. *Journal of the Academy of Marketing Science*, *45*(2), 135-155.

Mason, B. M., Mesaglio, T., Barratt Heitmann, J., Chandler, M., Chowdhury, S., Gorta, S. B. Z., Grattarola, F., Groom, Q., Hitchcock, C., Hoskins, L., Lowe, S. K., Marquis, M., Pernat, N., Shirey, V., Baasanmunkh, S., & Callaghan, C. T. (2025). iNaturalist accelerates biodiversity research. *BioScience*, *75*(11), 953–965.

Matthiopoulos, J., Field, C., & MacLeod, R. (2019). Predicting population change from models based on habitat availability and utilization. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1901), 20182911.

McIntosh, B. S., Ascough, J. C., Twery, M., Chew, J., Elmahdi, A., Haase, D., Harou, J. J., Hepting, D., Cuddy, S., Jakeman, A. J., Chen, S., Kassahun, A., Lautenbach, S., Matthews, K., Merritt, W., Quinn, N. W. T., Rodriguez-Roda, I., Sieber, S., Stavenga, M., . . . Voinov, A. (2011). Environmental decision support systems (EDSS) development – Challenges and best practices. *Environmental Modelling & Software*, *26*(12), 1389-1402.

Merkle, E. C., & Hartman, R. (2018). Weighted Brier score decompositions for topically heterogenous forecasting tournaments. *Judgment and Decision Making*, *13*(2), 185-201.

Metzger, F. M., & Runge, G. (2023). Reciprocity of Data Sharing Infrastructures: A Conceptual Norms Framework. In A. Roßnagel, M. Friedewald, C. L. Geminn, M. Karaboga, & S. Schindler (Eds.), *Data Sharing – Datenkapitalismus by Default?* https://doi.org/10.24406/publica-2072

Morrison, C. A., Robinson, R. A., Leech, D. I., Dadam, D., & Toms, M. P. (2014). Using citizen science to investigate the role of productivity in House Sparrow Passer domesticus population trends. *Bird Study*, *61*(1), 91-100.

Natural Capital Project. (2025). InVEST 3.16.2: Stanford University, University of Minnesota, Chinese Academy of Sciences, The Nature Conservancy, World Wildlife Fund, Stockholm Resilience Centre and the Royal Swedish Academy of Sciences. https://doi.org/10.60793/natcap-invest-3.16.2

Nghiem, L. T. P., Papworth, S. K., Lim, F. K. S., & Carrasco, L. R. (2016). Analysis of the Capacity of Google Trends to Measure Interest in Conservation Topics and the Role of Online News. *PLOS ONE*, *11*(3), e0152802.

ONS Geography. (2025). *Rural Urban Classification (2021) of Output Areas in EW*. Retrieved 2025-05-19 from https://geoportal.statistics.gov.uk/datasets/ons::rural-urban-classification-2021-of-output-areas-in-ew/about

Pai, P., & Tsai, H.-T. (2016). Reciprocity norms and information-sharing behavior in online consumption communities: An empirical investigation of antecedents and moderators. *Information & Management*, *53*(1), 38-52.

Peach, W. J., Mallord, J. W., Ockendon, N., Orsman, C. J., & Haines, W. G. (2015). Invertebrate prey availability limits reproductive success but not breeding population size in suburban House Sparrows Passer domesticus. *Ibis*, *157*(3), 601-613.

Pearce, J. M., & Parncutt, R. (2023). Quantifying Global Greenhouse Gas Emissions in Human Deaths to Guide Energy Policy. *Energies*, *16*(16).

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, *54*(4), 1643-1662.

PurpleAir. (2025). *PurpleAir*. Retrieved 2025-10-03 from https://purpleair.com/

Regos, A., Tapia, L., Gil-Carrera, A., & Domínguez, J. (2021). Caution Is Needed When Using Niche Models to Infer Changes in Species Abundance: The Case of Two Sympatric Raptor Populations. *Animals*, *11*(7), 2020.

RHS. (2025). *RHS State of Gardening Report 2025*. https://www.rhs.org.uk/about-us/what-we-do/rhs-state-of-gardening

Robinson, D. L., Goodman, N., & Vardoulakis, S. (2023). Five Years of Accurate PM2.5 Measurements Demonstrate the Value of Low-Cost PurpleAir Monitors in Areas Affected by Woodsmoke. *International Journal of Environmental Research and Public Health*, *20*(23).

Rue, H., Riebler, A., Sørbye, S. H., Illian, J. B., Simpson, D. P., & Lindgren, F. K. (2017). Bayesian Computing with INLA: A Review. *Annual Review of Statistics and Its Application*, *4*(Volume 4, 2017), 395-421.

Sandhaus, S., Kaufmann, D., & Ramirez-Andreotta, M. (2019). Public participation, trust and data sharing: gardens as hubs for citizen science and environmental health literacy efforts. *International Journal of Science Education, Part B*, *9*(1), 54-71.

Sharma, N., Greaves, S., Siddharthan, A., Anderson, H. B., Robinson, A.-M., Colucci-Gray, L., Wibowo, A. T., Bostock, H., Salisbury, A., & Roberts, S. (2019). From citizen science to citizen action: analysing the potential for a digital platform to cultivate attachments to nature. *Journal of Science Communication*, *18*(01), A07.

Shaw, L. M., Chamberlain, D., & Evans, M. (2008). The House Sparrow Passer domesticus in urban areas: reviewing a possible link between post-decline distribution and human socioeconomic status. *Journal of Ornithology*, *149*(3), 293-299.

Siegert, S. (2017). Simplifying and generalising Murphy's Brier score decomposition. *Quarterly Journal of the Royal Meteorological Society*, *143*(703), 1178-1183.

Teo, L. W. (2023). *House Sparrows dig and chew on grass stolons*. https://besgroup.org/2023/05/18/house-sparrows-dig-chew-grass-stolons/

Trantas, A., Plug, R., Pileggi, P., & Lazovik, E. (2023). Digital twin challenges in biodiversity modelling. *Ecological Informatics*, *78*, 102357.

Vehtari, A. (2025). *Can cross-validation be used for hierarchical / multilevel models?* https://users.aalto.fi/~ave/CV-FAQ.html#8_Can_cross-validation_be_used_for_hierarchical__multilevel_models

Vincent, K. (2005). *Investigating the causes of the decline of the urban house sparrow Passer domesticus in Britain* [PhD Thesis, De Montfort University, Leicester].

Walling, E., & Vaneeckhaute, C. (2020). Developing successful environmental decision support systems: Challenges and best practices. *Journal of Environmental Management*, *264*, 110513.

Weber, M. M., Stevens, R. D., Diniz-Filho, J. A. F., & Grelle, C. E. V. (2017). Is there a correlation between abundance and environmental suitability derived from ecological niche modelling? A meta-analysis. *Ecography*, *40*(7), 817-828.

Weir, J. E. S. (2015). *Urban green space management for invertebrates and house sparrows* [PhD Thesis, Imperial College London, Department of Ecology and Evolution].

Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. MIT University Press.

Williams, R. L., Stafford, R., & Goodenough, A. E. (2015). Biodiversity in urban gardens: Assessing the accuracy of citizen science data on garden hedgehogs. *Urban Ecosystems*, *18*(3), 819-833.