

Manuscript Number: LAND-D-14-00393R1

Title: Modelling biodiversity distribution in agricultural landscapes to support ecological network planning

Article Type: Research Paper

Keywords: citizen science; ditch; habitat connectivity; wetland; agricultural ditch; dispersal corridor

Corresponding Author: Dr. Hannah Louise Mossman, Ph.D

Corresponding Author's Institution: Manchester Metropolitan University

First Author: Hannah Louise Mossman, Ph.D

Order of Authors: Hannah Louise Mossman, Ph.D; Christopher J Panter, BSc; Paul M Dolman, PhD

Abstract: Strategic approaches to biodiversity conservation increasingly emphasise the restoration of ecological connectivity at landscape scales. However, understanding where these connecting elements should be placed in the landscape is critical if they are to provide both value for money and for biodiversity. For such planning to be effective, it is necessary to have information of the distributions of multiple taxa, however, this is of poor quality for many taxa. We show that sparse, non-systematically collected biological records can be modelled using readily available environmental variables to meaningfully predict potential biodiversity richness, including rare and threatened species, across a landscape. Using a large database of ad-hoc biological records (50 501 records of 502 species) we modelled the richness of wetland biodiversity across the Fens, a formerly extensive wetland, now agricultural landscape in eastern England. We used these models to predict those parts of the agricultural ditch network of greatest potential conservation value and compared this to current strategic network planning. Odonata distribution differed to that of other groups, indicating that single taxon groups may not be effective proxies for other priority biodiversity. Our results challenged previous assumptions that river channels should comprise the main connecting elements in the Fens region. Rather, areas of high ditch density close to a main river are likely to be of greater value and should be targeted for enhancement. This approach can be adopted elsewhere in order to improve the evidence-base for strategic networks plans, increasing their value for money.

Dear Editor

Thank you for the detailed and insightful comments from reviewers and we are grateful for the invitation to submit a revised version of the paper.

We have substantially revised our ms in order to address the reviewers comments, including more clearly stating our aims and objectives, clarifying aspects of our methodology and highlighting the wider relevance of the paper.

We have carefully considered Reviewer 2's comments on our modelling methodology and conducted a detailed review of the literature. Following this review, we have retained our original approach because we believe that it is the most appropriate given the challenges of our data set. We provide a detailed discussion of other methodological approaches and why we do not believe they are more suitable in this instance in response to Reviewer 2, and have included a shortened version of this in our revised ms.

We are grateful to the reviewers for their detailed and careful consideration of the manuscript. Revising the paper to address all comments of the reviewers has helped us to substantially improve the paper, which we are now confident is clearer and more robust.

Below, we list the original comments and beneath each detail *in italics* our response. We hope the paper will now be suitable for publication in Landscape and Urban Planning.

Best regards,
On behalf of all Authors
=====

Dear Editor

Thank you for the detailed and insightful comments from reviewers and we are grateful for the invitation to submit a revised version of the paper.

We have substantially revised our ms in order to address the reviewers comments, including more clearly stating our aims and objectives, clarifying aspects of our methodology and highlighting the wider relevance of the paper.

We have carefully considered Reviewer 2's comments on our modelling methodology and conducted a detailed review of the literature. Following this review, we have retained our original approach because we believe that it is the most appropriate given the challenges of our data set. We provide a detailed discussion of other methodological approaches and why we do not believe they are more suitable in this instance in response to Reviewer 2, and have included a shortened version of this in our revised ms.

We are grateful to the reviewers for their detailed and careful consideration of the manuscript. Revising the paper to address all comments of the reviewers has helped us to substantially improve the paper, which we are now confident is clearer and more robust.

Below, we list the original comments and beneath each detail *in italics* our response. We hope the paper will now be suitable for publication in Landscape and Urban Planning.

Best regards,
On behalf of all Authors
=====

REVIEWER 1

General comments

Overall this is a very useful study of fenland biodiversity and will help to formulate the 'fens for the future' strategy regarding connecting landscape features throughout the Fens. The Modelling strategy is an interesting one and appears to get round the sparse data issues that often prevail in areas of low biodiversity. It is of interest to an international audience in terms of its methods but its relevance beyond the fenland basin needs to be better explained in the discussion. The paper is well written and clear though note specific queries below.

We are pleased that this reviewer has such positive comments on the tangible utility of the analyses presented here, considers the modelling robust and the paper generally well written and clear. We have amended the discussion to highlight the interests to wider audiences. The paper describes a methodology for providing an evidence base to underlie conservation planning at landscape-scales, outlining an approach that is applicable to any region with a minimum level of biological recording; we have highlighted and strengthened this aspect of the discussion.

Detailed comments

Table 1 Environmental predictor variables are described in the text and listed in table 1 but there is no justification for the choice of these variables. ie why were these chosen and what would they be expected to predict in terms of the relationship between the variable and biodiversity, much is implicit rather than explicit? I would have expected ditch characteristics to be very important environmental variables, eg. overall cross-sectional dimensions , permanence or not of water through the year, water quality and ditch/ditch margin enhancement schemes in place or not. Would it be possible to refine this modelling exercise with at least some recognition that ditches vary greatly in biodiversity based on ditch characteristics not just distance/density indices as used here? Perhaps the IDBs could provide this information for the model? The authors do recognise this as an issue in lines 301-302 but it may be that these ditch characteristics are much more important than those used.

We agree with the reviewer that other ditch characteristics, such as water quality, ditch margin enhancements, and the depth and flow rates of water are likely to have important influences on ditch biodiversity. However, we believe a key strength of our approach is its application to the landscape scale using widely available environmental predictors. These coarse-scale predictors performed well, highlighting interesting and consistent patterns. A consequence of working at this basin-scale extent inevitably means that detailed field and local scale measures are not available. For example, water quality/chemistry measurements are made at water monitoring stations operated by the Environment Agency, but these are scantily distributed across our region and restricted to main rivers and the larger drains (<http://maps.environment-agency.gov.uk/wiyby/wiybyController?ep=maptopics&lang=e>). Some Internal Drainage Boards (IDBs) have data on water depth and flow rates, but many do not and where data do exist, they are restricted to the larger drainage channels. Some IDBs also have some information about the frequency of ditch management, but many do not hold this data or have it accessible, and ditch management may also be subject to ad hoc landowner management. Parallels would exist for many other widely distributed landscape elements in other landscape types or regions, for example in analysing biodiversity potential of grassland or hedgerows in agricultural landscapes it is unlikely that detailed data on structure and quality of individual hedgerows or grassland parcels will be available. Importantly, the approach outlined here shows how useful predictions can be made in the

absence of such field-based data. We have amended the ms, particularly in the introduction, to ensure that this aim is clear.

167-171 Explain this in more detail so that implications are clear.

We have now revised this ms to make this explicit.

204 'The Great fens' is not the correct term. I assume this refers to the Wildlife Trust's Great Fen Project?

The reviewer is correct - Great Fens Project is the restoration area surrounding the remnant wetlands of Holme and Woodwalton Fens. We apologise for this error and have amended the ms accordingly.

215-216 Not clear what this means exactly. Provide more detail.

The original statement added little to the ms and we have deleted it for brevity.

228 There needs to be some explanation of what is meant by the term 'Minimum adequate model'

'Minimum adequate model' is a standard statistical term that refers to a simplified model, constructed by removing non-significant variables, in order to provide a parsimonious explanation of the observations. However, we have replaced this with the more widely used term 'Minimum model', which is defined in the methods at lines 227-230.

279 Can you attempt to explain the low biodiversity of rivers flowing through silt soils?

295 the issue of silt soils arises again here but no explanation of causal mechanisms. I assume that the predictions referred to here are of low rather than high biodiversity?

We have edited to make explicit that silt soils are associated with lower levels of biodiversity. Our finding that ditch landscapes on silt soils have lower biodiversity potential is consistent across taxonomic groups. The ecological mechanisms underlying this negative association were not tested in this paper and so we can only speculate. Soils with finer particles, such as silts, tend to contain higher nutrient concentrations and so the negative association between species richness and silt soils may reflect reduced water quality. We have now more fully discussed potential reasons for this finding (lines 343-353)..

315 it would be worth reflecting on the relative mobility of the odonata compared to many other wetland etc species with regard to network planning.

In addition to Odonata, many other wetland invertebrates are also capable of active dispersal over longer distances, such as some groups of water beetles. The dispersal distances of more weakly-flying species may be increased by the wind. Other species, including plants, have adaptations for aerial or vertebrate-assisted dispersal, or have intraspecific variations in dispersal ability that allow episodic dispersal events (see Bilton et al (2001) Ann. Rev. Ecol & Syst). Aquatic plants are capable of dispersal of vegetative propagules (e.g. root and rhizome fragments) via waterways. We therefore do not think it would be helpful to single out Odonata as being particularly good dispersers.

REVIEWER 2

This paper analyses unstructured species occurrence records in the English Fens in relations to a number of explanatory variables. The resulting extrapolated species richness surfaces are then used to inform connectivity planning in the region.

The paper is well written and the study goals are interesting and worthwhile. However, I had some serious concerns about the methodology used and unless the method can be adequately validated I did not feel it worthwhile discussing any finer details.

We are pleased the reviewer considers the paper well written and potentially worthwhile. We have addressed in detail below the methodological issues.

Species richness

Firstly, there is a substantial literature on problems and potential solutions to control for spatio-temporal variation in recorder effort in the analysis on unstructured species occurrence data (Hassall & Thompson 2010; Kéry et al. 2010; Hill 2011; Wintle et al. 2012; van Strien et al. 2013; to highlight just a few). This literature is effectively ignored in this paper and instead a very basic method is used which is unlikely to adequately control for recorder effort.

The authors state: "In order to account for recorder effort, the square root transformed number of records of a given group of species was included in the models of richness as a covariate". This assumes a specific Power law relationship between sampling effort and number of species. However, where is the evidence that such a relationship is correct? The authors need to validate their method. If the solution to the huge biases from recorder effort were so simple, why the need for the extensive literature on the subject (e.g. benchmarking methods, occupancy modelling).

The authors simply compare their method with two very simplistic methods 1. Excluding squares with few records (e.g. the lowest quartile of records)- which assumes the problem doesn't exist at thresholds higher than this, or 2. weighting models by the number of records, or including number of records as an untransformed covariate- which assumes a clearly incorrect linear relationship between sampling effort and number of species. In addition, it is not very clear how these different methods are compared. The authors state "predictive performance of models calibrated with 75% of the data was assessed across 100 bootstrap samples", so is this test using the same data (for which species richness is biased by recording effort?). I'm unclear how this validates the method to show it gives accurate measures of 'true' species richness.

We have undertaken a substantial review of relevant literature and, while we agree we need to refer to this literature and justify our methods in relation to it, we have not found any method that would be applicable to the varied challenges posed by our dataset. We have however, provided a fuller explanation of these issues, and the reasons for the analytical approach adopted, in the methods of the revised paper (lines 200-220, 382-395). We also provided a detailed description of the issues below.

The failure of recorders to detect all species has three main consequences for models (Kéry et al. 2010 J Biogeography) – (1) models underestimate species richness, (2) covariates are biased towards zero and (3) variables that are correlated with recording effort may be wrongly identified as influencing species richness. We believe our method (including the number of records as a covariate to statistically control for spatial variation in recorder effort) deals with the latter, but agree with the reviewer that the first two problems are likely to apply. The first two issues affect the accuracy of absolute estimates of species richness, while the third issue affects whether spurious predictor variables are selected and whether predicted species richness provides an index of true species richness. Our aims are to

determine patterns in relative (not absolute) species richness and identify important environmental variables, therefore issue three is by far the most important for our purposes.

Hierarchical occupancy models use repeat visits to a site to allow the detection process to be modelled as well as processes influencing occurrence/ abundance – thus potentially dealing with issues 1 and 2. Our data make the application of these methods untenable. Firstly, we are dealing with a very large number of species, many of which are rare. Secondly, the underlying state of our system (i.e. the true occurrence status of species i at site j) cannot be assumed to be constant as species may have colonised or become locally extinct. Thirdly, it is uncertain how to define a visit – it could be arguably defined as unique recorder-date combinations, or restricted to recorder-date combinations where one of the target species was detected (thus meaning the recorder was looking for a species – but what if they did not see any despite looking?), or where a recorder has recorded a species at some point (so has the potential to be looking for a species). In our data set, the threshold for defining a visit has a big effect on the number of visits per site (sometimes an order of magnitude), so presumably will have a big effect on estimated detection probabilities. Fourthly, our site by visit by species matrix is much larger than applications of occupancy modelling that we are aware of, and this matrix is dominated by zeros. Authors have been forced to remove rare species due to issues with model convergence (e.g. Gilroy 2014 J Applied Ecology) – these rare species are very important for this study so removing them would have a serious effect on our results. We note that extensions are available for some of these individual issues (e.g. extensions that allow the state of the system to vary, or that allow for multiple species), however, we believe that simultaneously dealing with these issues would require substantial statistical development beyond the state-of-the art. We therefore consider that hierarchical occupancy models, in their current state of development, are not appropriate for our dataset.

Benchmark species (a species or set of species assumed to be widespread so that their recorded occurrence is assumed to relate to recording effort rather than environmental conditions) can provide a useful proxy for recorder effort (e.g. Hill 2011), so can potentially help deal with issue 3. A good benchmark species can be assumed to be ubiquitous. However, we expect the distributions of wetland priority taxa to be strongly patchy due to environmental conditions, so these species would be poor benchmarks. Additionally, we have direct data on recording effort (number of records), and feel it is preferable to use this rather than a proxy.

We do not feel the general reader would benefit particularly from a detailed explanation for not adopting each of these alternative analytical approaches, but have nevertheless summarised the points above in the methods text.

The rationale for using the number of records as a covariate to account for spatial variation in recording effort is consistent with species rarefaction / accumulation curves. We could have used rarefaction to estimate the species richness of each square for a set number of records, however this would have been likely to overestimate the effect of recorder effort on species richness (as it assumes that observed species richness is primarily due to recorder effort, with the residuals subsequently explained by environmental variables, see Freckleton 2002 J Animal Ecology for critique of such regression of residuals approaches). By including the number of records as a covariate in GLMs we could simultaneously parameterise the effect of recorder effort and environmental variation on observed species richness. Based on the form of species accumulation curves, we would expect a saturating or power law relationship between number of records and proportion of total species richness that was recorded. We agree that the assumption that this forms a specific power law relationship was not justified, so re-ran our models with a polynomial term for recorder effort. However,

the existing model (with square root the number of records) explained a greater proportion of deviance, so was retained.

Including the number of records as a covariate allows us to evaluate the conditional effect of environmental variation on species richness accounting for spatial variation in recorder effort. This means we avoid the issue 3, where variables that are correlated with recorder effort are wrongly identified as influencing species richness. Methodological developments (e.g. occupancy modelling) now enable issues 1 and 2 to be addressed, but extending these methods for our dataset would require substantial, non-trivial methodological development. As issues 1 and 2 are less serious for our study, we have continued with our existing methods, but have added the caveat that our estimates of species richness are best thought of as an index of relative richness rather than absolute measures of true richness. We have also added more discussion justifying our methods and explaining why other existing alternative methods were not suitable for our purposes.

We note the issues with our cross-validation approach raised by the referee, so have removed this from the methods. We agree that in such a mega- but sparse-data situation, any analytical framework will have limitations, but results are clear and consistent with ecological understanding. If we have missed any relevant methods that would be both appropriate and practicable, we would of course be happy to consider using them.

Connectivity modelling

The authors use their models to produce maps of species richness which are then used to inform the best places to put corridors to link high biodiversity areas, comparing their suggestions with an action plan collated primarily from expert opinion. Although the aim here is laudable, I am not sure about the logic behind the approach. In addition to the fact that species richness maps produced in this paper stem from data from 1987 onwards (and so species richness may be very different now) and the questions raised above about the appropriateness of the methods to actually produce species richness maps, I could not quite understand some of the reasoning used in producing the potential corridors. Unless I am mistaken, the criteria that the authors use to define corridors is that 1) they must connect areas of known high biodiversity (which makes sense) and 2) the corridors must occur in areas of high predicted biodiversity. My understanding is that a conservation corridor involves restoration of key areas of hostile habitat in order to link isolated populations. If the populations are already linked by high quality habitat do they need a corridor? Or are they in fact already linked?!

First, our models do not predict areas of high current biodiversity, rather they predict areas that have the potential for high biodiversity richness, based on their soil and other landscape variables, and we have linked these together with our corridor. However, we do not know if the cells of our predicted corridor currently realise that biodiversity potential with their existing habitat (which could be improved by enlarging, and sympathetic management to restore and enhance habitat quality), or if they have poor current habitat but high potential on the basis of e.g. appropriate soil type, ditch density and proximity to main channel etc. In either case we predict the potential to enhance the biodiversity value of those areas to be greater than in other parts of the landscape that have low potential for biodiversity and thus we are recommending these areas should be targeted for enhancement. Throughout the revised manuscript, we have edited to emphasise this rationale for focusing on biodiversity potential.

Second, the use of records since 1987 was a compromise between reflecting the current or recent distribution of species and including sufficient records in the dataset to capture the rare species that are infrequently recorded and to capture data from sparsely visited parts of the landscape. There may have been local extinctions since 1987, but it is

likely that the landscape-scale predictors we consider here remain suitable for those species and there is therefore the potential for recolonisation. We have amended the ms to clarify this (lines 96-99).

REVIEWER 3

Reviewer #3: Major comments:

No!

We are pleased that the reviewer has not identified any issues requiring major attention.

Minor comments:

L 58-64 Add your hypothesis or research questions

We amended the ms to now include a clear statement of our aims (lines 66-69).

L 68, figure 1. I recommend to leave "a" part black and white to the main text. Make map "b" colourful and add it to supplementary material or leave to the main text but as the colourful map. At the moment, it does not tell much about number of records.

This figure has been revised accordingly.

Figure 1. Map ranges (no of records) are not equal! I think it is misleading and recommend to make all ranges equal!

L255/figure 3. Once more: ranges. Per each biological variables the ranges should be equal.

The number of records is highly skewed, with many values of less than 50, but a few extreme values of three orders of magnitude higher. With such skewed data, a map using equally sized class intervals would not be informative, because it would require either hundreds of class (most of which would not contain a value) or all the squares to be in the same class with the exception of three or four extremely high values. We have therefore allocated divisions ('bins') using Jenks natural breaks classification, which one of the standard methods in GIS for determining class interval breaks with such data.

L 76-89. My concerns about biological data. Firstly, you should explain in this paragraph, how reliable this data are. How much we can trust this data? In my experience, not all "citizen science" collected data are not trustable. Did you check it and omitted something? This all should be stated in this paragraph.

The reviewer raises an important issue that we had initially omitted for brevity. We obtained records from Biological Records Centres and natural history societies who verify and validate the records they receive with local taxonomic experts (usually county recorders). Some National Biodiversity Network (NBN) data are not validated at source, however, we further validated our collated species lists with the help of local taxonomic experts. We thus have strong confidence in the reliability of our data and have modified the ms to include the validation process (lines 89-92).

L 76-89. Secondly, how did you managed to put all this data to 1-km squares. Did you really have all available data connected with exact geographical coordinates? Explain it all in the paper.

We have now clarified within the paper (lines 99-107), now including that the majority (74%) of records we initially collated were resolved to a spatial resolution of 1 km or finer and we aggregated these records to the scale of 1-km grid cells. A small proportion of plant records were recorded at the tetrad scale and were assigned to all of the four 1-km squares comprising the tetrad. The small number of records at coarser spatial resolutions were excluded.

L 85 Please write more clearly what was exact "study period". I assume 1987-2014, is it correct? Put full time period somewhere in methods.

We have amended the methodology to clearly state our study period, 1987-2012 inclusive (lines 83-84, 93-95). We have also included additional text to clarify the consequences of selecting this time period.

L 193 Did you use variance partitioning?

We did use variance partitioning and have amended the text to clarify this.

L 173-194 What statistical program and/or packages did you use?

Analyses were conducted in the statistical package in R. We have now referenced R in the methods.

L 240/figure 2. Abbreviations are actually missing. You should put them to the figure 2 legend or table 1 or in the main text. At the moment, figure 2 are not self-readable because I cannot understand all those abbreviations.

We have included abbreviations in Table 1 and have referred to them in the legend of Fig. 2.

TITLE PAGE

Modelling biodiversity distribution in agricultural landscapes to support ecological network planning

Hannah L MOSSMAN*

Christopher J PANTER

Paul M DOLMAN

School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ

* Current address: Division of Biology and Conservation Ecology, School of Science and the Environment, Manchester Metropolitan University, Chester St, Manchester, M1 5GD.

Corresponding Author:

Hannah L Mossman, Division of Biology and Conservation Ecology, School of Science and the Environment, Manchester Metropolitan University, Chester St, Manchester, M1 5GD. Tel: +44 (0) 161 247 1186. Email: h.mossman@mmu.ac.uk

Co-author email addresses:

Christopher J Panter: c.panter@footprint-ecology.co.uk

Paul M Dolman: p.dolman@uea.ac.uk

HIGHLIGHTS

- We used ad-hoc biological data to model landscape-scale wetland species richness.
- Models were used to assess and improve a proposed ecological connectivity network.
- Our evidence-based network was shorter and connected areas of higher richness.
- Our results challenge previous assumptions of important network elements.
- Odonata were poor proxies for other groups of wetland species.

1 MANUSCRIPT

2 **1. Introduction**

3

4 Habitats are increasingly fragmented. Furthermore, in human landscapes, habitat patches are
5 often surrounded by land uses that are potentially hostile to dispersal, increasing functional
6 isolation (Nowicki et al., 2014). Such habitat fragmentation and isolation increase local
7 population vulnerability to extinction and reduced dispersal opportunities limit species'
8 ability to respond to climate change, further reducing biodiversity resilience (Hill et al.,
9 2002). Strategic approaches to conservation are, therefore, increasingly focused at the
10 restoration of landscape connectivity by the creation of movement corridors, stepping stones
11 or by improving landscape permeability (Dolman, 2012; Lawson et al., 2012; Saura et al.,
12 2014). However, the nature, size and placement of these connecting elements are critical if
13 investment of finite funds and land resources are to give optimal returns. There are several
14 key issues to the success of landscape connectivity; identifying what species should be
15 targeted within a landscape (Dolman et al., 2012), ensuring that the connectivity elements
16 comprise habitats that suit these species and establishing where these connecting elements
17 should be placed.

18

19 Ecological networks are often designed to enhance the metapopulation viability of individual
20 high profile species, such as top predators (Klar et al., 2012) or other mobile species (Bani et
21 al., 2002). However, the ability of such species to act as connectivity umbrellas for
22 assemblages of other species may be limited (Cushman & Landguth, 2012) because the
23 suitability of the habitat and type of connecting element differs amongst taxa. For example,
24 while linear field margins may provide connectivity to some generalist butterflies (Delattre et

25 al., 2010), they may act as sinks to other taxa (Krewenka et al., 2011). Similarly, hedgerows
26 are often purported to provide suitable corridors for woodland species, but may only provide
27 habitat for woodland edge species (Liira & Paal, 2013). The planning of landscapes to
28 provide resilience for assemblages of regional biodiversity therefore requires the
29 consideration of multiple relevant taxa (Zulka et al., 2014).

30

31 Decisions regarding the optimum placement of connecting elements should be made using
32 evidence of the current and potential distribution of a full complement of target species.
33 Existing protected sites that retain a concentration of rare species generally form the focus of
34 connectivity networks (Beier et al., 2011) and the existence of species within these fragments
35 is often well known. However, our understanding of the distribution of species throughout the
36 rest of the landscape is incomplete, with some locations receiving high levels of recording
37 effort and others very little. Poorly recorded areas that are nevertheless potentially suitable
38 for a species may harbour unrecorded residual populations, or be more likely to be colonised
39 if both habitat quality and connectivity are improved (Lawson et al., 2014). Unsystematically
40 collected biological data therefore do not provide a reliable assessment of conservation value
41 or potential across a landscape. This results in reliance on expert opinion in the design of
42 landscape connectivity (Beier et al., 2009; Eycott et al., 2011). However, if the patchy nature
43 of recording effort is accounted for in the analysis (Kéry, 2011), ad hoc biological data can be
44 exploited to provide more objective design of landscape connectivity.

45

46 In this study, we use the Fens, a formerly extensive wetland system in eastern England, to
47 demonstrate how connectivity planning can be informed by modelling ad hoc biological
48 records with easily obtainable, landscape-scale environmental data. Remaining wetland

49 habitat in the Fens is highly fragmented and isolated within an intensive agricultural
50 landscape, but there is a high potential for connectivity through enhancing management of
51 linear drainage ditches. Ditches in intense agricultural areas are often rather different to
52 natural streams (Herzon & Helenius, 2008), supporting lower biodiversity (Williams et al.,
53 2004); however, they can act as reservoirs for important regional wetland biodiversity (Simon
54 & Travis, 2011). Biological recording within the wider Fens landscape is extremely sparse, so
55 simple mapped biological richness cannot be used as an evidence base for selecting
56 potentially biodiverse ditches for improved management or in the design of connectivity
57 networks. Recent attempts at strategic planning (e.g. Fens for the Future Partnership, 2012)
58 have therefore relied on a combination of expert opinion and untested assumptions of where
59 this targeted management should be placed.

60

61 We take the approach of modelling potential biodiversity value in relation to underlying
62 environmental factors and landscape context, to predict where in the landscape targeted
63 management to enhance habitat quality will have greatest potential to support biodiversity
64 and enhance connectivity. We use an extensive but unevenly distributed database of 67,395
65 ad hoc biological records to model the richness of groups of wetland species across the Fens
66 landscape in relation to a range of coarse-scale environmental and landscape factors. Using
67 these models, we aim to: 1) predict and map the potential richness of groups of wetland
68 species in order to identify parts of the landscape of greatest potential conservation value; 2)
69 apply these maps of predicted biodiversity potential to assess current strategic planning maps.

70

71 **2. Methodology**

72 *2.1 Study area*

73 The Fens, covering almost 4,000 km² of eastern England (Fig. 1), was formerly an extensive
74 wetland area but only 1% of wetland habitat remains. This habitat is concentrated in six key
75 protected areas, which are each small (mean area 819 ha) and isolated within the country's
76 most important arable agricultural landscape. More than 20 million km of ditches and
77 drainage channels criss-cross the Fens landscape and by targeting selected ditches for
78 enhanced management, the ditch network presents an excellent opportunity for increasing
79 both habitat area and connectivity for wetland species. However, the current conservation
80 value of large parts of this landscape is poorly known.

81

82 *2.2 Biological data*

83 All available species observations (records) were collated for the period of 1987-2012 from
84 the 4147 1-km squares wholly or partly within the Fens Natural Character Area boundary
85 (Natural England, 2013), with an extension (3 km from the boundary) to include Chippenham
86 Fen, one of the three important relict fen sites in the Fens. Records were compiled from Local
87 Biological Records Centres, the National Biodiversity Network (NBN) gateway, national and
88 county natural history and recording societies whose records were not available via NBN, and
89 unpublished documents or reports. Records sent to Biological Records Centres and societies
90 are validated by expert county recorders. Although NBN data may include some unvalidated
91 records submitted by the public, our collated species lists were validated by a range of local
92 taxonomic experts. Records were managed using RECORDER 6 software (Joint Committee
93 for Nature Conservation, Peterborough, UK). The study period (1987-2012 inclusive) was
94 selected as a compromise between reflecting the current or recent distribution of wetland
95 species and including sufficient records in the dataset to capture rare species and the potential
96 distribution of sparsely recorded taxonomic groups. There may have been local extinctions
97 since 1987 due to local changes in habitat quality, nevertheless the landscape predictors we

98 consider will indicate the biodiversity potential should habitat and connectivity be restored.
99 The majority (74%) of records were resolved to a spatial resolution of 1 km or finer and these
100 were aggregated and analysed at the scale of 1-km grid cells. Tetrad records were assigned to
101 all of the four 1-km squares comprising the tetrad; species records at coarser spatial
102 resolutions were excluded. A small number of records of taxa not recorded to species level
103 were removed. Records of marine species were excluded, but those tolerant of brackish
104 conditions were retained. Following additional filters described below (e.g. removal of
105 coastal squares), a database of 255 291 records remained, of which 50 501 were records of
106 wetland plants (including conservation priority species) and conservation priority wetland
107 invertebrates. Conservation priority species were recognised as those designated as UK
108 Biodiversity Action Plan, Global and UK Red Data Book (except Least Concern), Nationally
109 Rare, Nationally Scarce or Nationally Notable A and B, according to JNCC (2012), plus
110 undesignated species with >25% of their UK distribution occurring in the Fens region –
111 hereafter referred to as ‘Fens Specialists’.

112

113 The richness of groups of wetland species were used as the biological response variables.
114 Seven widely recorded groups of wetland species were selected for modelling that were
115 considered good indicators of ditch quality: all Odonata (dragonflies and damselflies, 28
116 species), wetland plants (212 taxa), fully aquatic plants (137 taxa) and conservation priority
117 species (including plants and invertebrates) dependent on aquatic (fully aquatic and
118 submerged aquatic habitats, 90 species), littoral (aquatic margins, 109 species) or wetland
119 (208 species) habitats, and Fen Specialists (58 species). Wetland plant species were defined
120 as all Characeae (stoneworts, multi-cellular branched macro-algae) and those vascular plant
121 species associated with freshwater (aquatic, wetland or seasonally wet) habitats selected from
122 Hill et al. (2004) with Ellenberg moisture values ≥ 7 (species with Ellenberg salinity values of

123 >5 were excluded). Aquatic plants were a sub-set of the wetland plants, classified with
124 reference to existing lists by Palmer et al. (2013) and Mountford and Arnold (2006). The
125 autoecological requirements of conservation priority species and their association with
126 wetland, aquatic and littoral habitats, were classified following Mossman et al. (2012) and
127 Dolman et al. (2012).

128

129 *2.3 Environmental predictors*

130 The aim of this analysis was to predict the distribution of wetland species across the drainage
131 ditch network of the arable landscape based on readily-available, coarse-scale environmental
132 variables. Wetland Sites of Special Scientific Interest (SSSIs) were considered to be
133 reservoirs and potential sources of high quality biodiversity, therefore 1-km squares including
134 any part of a wetland SSSI were excluded from modelling. Wetland SSSIs were identified
135 based on the SSSI citation description (available at www.sssi.naturalengland.org.uk), with
136 wetland habitats considered to include ponds, gravel pits, wet woodland or carr, fen, bog,
137 grazing marsh and wet common.

138

139 Seventeen environmental predictors were initially selected as candidates for modelling (Table
140 1) based on ready availability across the study landscape and considered, *a priori*, to
141 potentially influence ditch biodiversity. A single value of each variable was calculated for
142 each 1 km square. The mean elevation above sea level, presence of an A or B road and the
143 distance from the centre of each 1-km square to the nearest wetland SSSI, Fenland island and
144 the edge of the Fen basin were calculated. Previous work has suggested that ditches with
145 highest conservation value are located near to the edge of the Fen basin or close to Fen
146 islands (Mountford & Arnold, 2006); the reasons for this are unclear, but may relate to high

147 water quality. Fen islands were delimited as areas of $>0.1 \text{ km}^2$ with an elevation of $\geq 5 \text{ m}$, and
148 the Fen basin defined as the 5 m contour boundary.

149

150 The soils of the Fens area are dominated by silt and peaty soil types. The percentages of each
151 1-km square comprising silt and selected peat soil types (Table 1) were calculated. Ditch
152 isolation from main channels and from tidal influence were considered potentially important
153 determinants of water quality, saline influence and thus of biodiversity richness. We
154 calculated the shortest network distance along the ditches and rivers network (extracted from
155 the Ordnance Survey (OS) surface water polylines, converted into a raster of 35m cells),
156 from the centre of each 1-km square to the nearest main channel/river and to the tidal
157 boundary, calculated in ArcGIS Spatial Analyst tools. Network distances were not weighted
158 by ditch size or type, such that all cells were assigned a value of 1. A cell size of 35 m was
159 sufficient to connect any small breaks in the polylines due to mapping error or underground
160 drains, but was considered small enough to prevent falsely connecting ditches in close
161 proximity that are not connected through surface water drainage. Some manual connections
162 were imposed on the network due to large breaks in the mapped surface, for example due to
163 bridges or pumping stations. Ditch density in each 1-km was calculated from OS polylines,
164 which defines both banks of ditches wider than 2 m; since ditches of $<2 \text{ m}$ in width are only
165 defined with one polyline, ditch density is an index that reflect both linear length and ditch
166 area.

167

168 The grades of the Agriculture Land Classification were used as proxies for potential
169 agricultural productivity, land-use intensity and therefore quality of both water and
170 banksides; this is an ordinal scale (1-5) where grade 1 is best agricultural land. The combined

171 percentage cover of grades 3 and 4, comprising the lowest quality agricultural land and
172 therefore representing the lowest intensity of agricultural land-use (no land was classified as
173 grade 5 in our study region), was used as a candidate predictor. The dominant land use in the
174 Fens region is arable; the percentage of each 1-km square comprising un-intensively managed
175 grassland (defined from Land Cover Map (Morton et al., 2011)) classes of Rough/Neutral
176 Grassland) was therefore considered of interest. The percentage of urban land use was also
177 calculated from OS data.

178

179 Inter-correlation among predictor variables was investigated using Pearson's correlation
180 coefficient and considered large enough to potentially have an effect on the models if $r > 0.5$,
181 following Freckleton (2002). Distance to the Fen basin was strongly correlated with distance
182 to the nearest wetland SSSI ($r=0.533$), network distance to the tidal boundary ($r=-0.523$) and
183 percentage of silt soils ($r=0.536$). Distance to the Fen basin was therefore excluded from the
184 modelling, whilst the other variables were retained.

185

186 Due to comprehensive county flora, plant species recording effort was substantially greater in
187 Norfolk and Suffolk relative to other counties. Therefore, to avoid spurious identification of
188 any environmental factor that differed between these and other counties, when modelling the
189 response of wetland and aquatic plant variables to environmental and landscape context
190 indicators, we included the two county groups as a binary covariate (0 = no flora, 1 = flora).

191

192 A number of 1-km squares were excluded from the models because they contained no surface
193 water, the surface water was more than 70 m from the nearest surface water feature (thus
194 indicating the feature was likely to be a pond rather than a ditch, contained part of a wetland

195 SSSI, or comprised >50% coastal area (defined using the Wash SSSI). This resulted in 3,745
196 1-km squares being used in analyses.

197

198 *2.4 Model construction*

199 *2.4.1 Accounting for recording effort*

200 It is well known that not all species present at a site will be detected and that this poses
201 challenges for analysis (Chen et al., 2013), as species richness is underestimated and
202 coefficients with environmental variables are closer to zero. Spatial variation in recorder
203 effort can have severe consequences for models, as environmental variables that are
204 correlated with recording effort may be spuriously identified as being related to species
205 richness. Hierarchical occupancy modelling can address these problems by utilising repeated
206 visits to the same site to estimate detection probabilities (MacKenzie & Kendall, 2002) and
207 thus has applications for analysing citizen science data (Isaac et al., 2014). Despite extensions
208 to deal with multiple species (Dorazio & Royle, 2005), application to datasets such as ours is
209 challenged by, for example, uncertainty in defining what represents a discrete ‘visit’, and
210 absence of information on visits that did not contribute species records to the data. An
211 alternative approach to addressing spatial variation in recorder effort is to include a proxy for
212 recorder effort as a covariate (Hill, 2011), allowing the conditional effects of environmental
213 variables on species richness to be assessed while controlling for recorder effort. We use the
214 total number of records in a 1-km square (i.e. including non-wetland species) as a proxy for
215 recording effort. We expect this relationship to be saturating as species accumulation curves
216 tend to saturate at high numbers of species, so we explored models using either square root
217 number of records or a polynomial term for number of records, using the former as it
218 explained more deviance. Although our method accounts for spatial variation in recorder

219 effort, we are unable to estimate the probability of *not* detecting a species, so our estimates of
220 species richness should be taken as an index of relative richness.

221

222 *2.4.2 Predicting species richness*

223 Statistical analyses were performed using the computing environment R (R Core Team,
224 2012). Predictor variables were standardised prior to modelling, with the exception of the
225 number of records. For each response variable, we fitted generalised linear models, with a
226 quasi-poisson error structure to deal with over-dispersion, containing all 16 predictor
227 variables (17 for wetland and aquatic plants owing to the inclusion of county). The full model
228 was simplified by backward elimination, judging variable retention by the t-test of β
229 estimates, with a threshold of $\alpha < 0.05$. The resulting minimum models were used to predict
230 the richness of each of the seven wetland species groups in each 1-km square of the study
231 area, with recording effort standardised as the overall median (41 records per 1-km square).
232 For the wetland and aquatic plant response variables, we standardised for the presence of a
233 recent flora by setting the value for all squares as 1.

234

235 Following Legendre and Legendre (2012), we used variance partitioning to calculate the
236 proportion of total variation in species richness explained by recording effort (total number of
237 records) and by environmental variables. To do this, we constructed models including 1) only
238 environmental conditions, 2) only recording effort and 3) both environmental conditions and
239 recording effort.

240

241 *2.5 Comparison of predicted biodiversity richness to the current strategic planning maps*

242 The 1-km squares were ranked by the predicted species richness for each of the seven
243 biological response variables separately, where a high rank (low number) was given to
244 squares with high predicted biodiversity. The mean of these ranks was calculated and
245 mapped. The resulting map of predicted biodiversity was compared to the Fens for the Future
246 Partnership (FFFP) (2012) strategic connectivity plan. The strategic connectivity network
247 consisted of three types of corridors: primary, secondary and landscape (Fens for the Future
248 Partnership, 2012). The primary corridor was the priority corridor and aimed to connect three
249 core areas thought to have high biodiversity value, the southern Fens and Ouse Washes,
250 Holme and Woodwalton fens (and associated Great Fens Project restoration area of the
251 Wildlife Trusts), and the Nene Washes. Secondary and landscape corridors aimed to provide
252 additional landscape connectivity; for the purposes of this study, secondary and landscape
253 corridors were combined.

254

255 We designed a new connectivity network that met with the objectives of the strategic
256 connectivity network and the following criteria. Corridors must connect areas of known high
257 biodiversity richness (wetland SSSIs) and presumed high richness, defined as those wetland
258 Local Wildlife Sites (LWS) that were $\geq 0.25 \text{ km}^2$ and occurred in areas of high predicted
259 biodiversity (richest $\geq 50\%$ of 1-km squares). A single primary corridor was placed to connect
260 the three core sites identified by the FFFP (2012). All corridors must join to form a
261 continuous network across the region and, where possible, achieve such connectivity by
262 passing through areas of greater predicted biodiversity.

263

264 The potential conservation effectiveness of the original strategic plan was compared to that of
265 the corridor network we proposed on the basis of the predicted distribution of wetland

266 biodiversity richness. These were assessed for each corridor strata (primary, secondary) in
267 terms of the length within each quartile of predicted species richness (for each 1-km square,
268 the mean of ranked richness across all the seven species groups). Proposed networks were
269 deemed to be more effective if a greater proportion of the corridors lay within the quartiles
270 predicted to be the most species-rich.

271

272 **3. Results**

273

274 *3.1 Effect of the environment on wetland biodiversity richness*

275 Overall, the minimum models explained 27.2 – 63.9% (mean = 40.3%) of the variation in
276 species richness of the seven groups (Table 2), performing best in predicting the richness of
277 wetland plants and aquatic plants (63.9% and 59.8%, respectively). A substantial part of the
278 explained variance was attributed to the independent effect of recorder effort (27.3 – 76.2%).
279 However, 17.1 – 52.8% of explained variance was attributed to the independent effect of
280 environmental variables, and a further 2.5-26.4 % to the joint effect of recorder effort and
281 environmental variables (Table 2). Species groups with the highest proportion of variance
282 explained by the environmental variables were Odonata, aquatic species and littoral species
283 (53%, 36% and 33% respectively).

284

285 The effects of many environmental predictors were consistent among species groups. Mean
286 elevation above sea level and percentage of urban area were not significant predictors of the
287 richness of any group (Fig. 2). A greater percentage of silt soil was negatively related to
288 species richness of all groups, compared to all types of peat soil (Fig. 2), although the

289 richness of wetland and littoral species were also lower with a greater percentage of deep
290 sand over peat or peat (Fig. 2).

291

292 Richness of all groups, except Fen Specialists, was greater closer to existing wetland SSSIs.
293 The richness of all groups except wetland plants, increased significantly with increasing
294 values of the index of ditch density (Fig. 2). The percentage of grade 3 and 4 agricultural land
295 (i.e. lower land-use intensity) was positively related to richness of Odonata, but not
296 significantly related to the richness of other groups. The richness of Odonata also increased
297 further from the tidal boundary; in contrast, the richness of aquatic species, and wetland and
298 aquatic plants was higher closer to the tidal boundary. The distance to a main river was not
299 significantly related to the richness of Fen Specialists and littoral species. Richness of the
300 remaining groups was highest closer to a main river, although predicted richness generally
301 decreased when main rivers were located on silt soils or were further from a wetland SSSI
302 (Fig. 3). The predicted richness of all groups was low around the coast (Fig. 3).

303

304 *3.2 Biodiversity potential of the proposed network corridors*

305 The combined predicted richness of ditch species suggests that the corridors of the proposed
306 strategic network are generally well placed (Fig. 4, 5). However, comparison of the strategic
307 map and the predicted biodiversity richness indicated that proposed corridors do pass
308 through some areas of lower biodiversity potential (Fig. 4). In contrast, our suggested map
309 achieved a greater proportion of connectivity in areas of high predicted richness (88% of our
310 corridors were located in the richest 50% of squares, compared to 66% of the FFTF corridors)
311 for a shorter overall length (27% shorter, combined primary and secondary corridors) (Fig.
312 5).

313

314 **4. Discussion**

315

316 Landscape connectivity and conservation plans are often developed with a reliance on
317 environmental and land cover data (Brooks et al., 2004a), but such broad data can be poor
318 surrogates for biodiversity (Araujo et al., 2001; Schindler et al., 2013), particularly for rare or
319 specialist species (Lombard et al., 2003). Ecological planning should consider the identity,
320 distribution and requirements of target species in that region, rather than being based on
321 untested assumptions of where species occur (Brooks et al., 2004b) as such assumptions can
322 lead to inappropriate selection of habitat type or placement of the connecting elements. For
323 example, the previous landscape connectivity plan in the Fens that was based on expert
324 opinion selected the main river channels as a key connecting component (FFTP 2012).
325 Whilst we found that species richness was higher closer to main river channels, rivers
326 flowing through areas of silt soils had particularly low predicted species richness, so
327 improvements to management or connectivity in these areas may have limited benefits for
328 wetland biodiversity. This has important implications for other landscapes where a single
329 land cover variable has been the focus of network planning, because without validating with
330 biological data the use of single features can prevent selection of optimal connectivity.

331

332 Increasing ditch density was a significant predictor of species richness for all groups, except
333 wetland plants. The ditch density was a particularly strong predictor of priority species (those
334 with a conservation designation) associated with littoral margins. Littoral species are of
335 particular conservation importance in the Fens region, but are often overlooked by
336 conservation interventions compared to submerged aquatic species (Mossman et al. 2012).

337 Thus specifically targeting areas of high ditch density close to rivers for improved
338 management, rather than the main river channels themselves, would substantially add
339 conservation value. This highlights the importance of considering the identity and
340 requirements of the species that are the priorities for conservation and connectivity in a
341 region or a landscape.

342

343 Several broad and readily available landscape variables, such as distance to a protected site
344 (SSSI) and cover of silt soils, were important predictors of biodiversity. Thus, such variables
345 can be used to select areas for restoration or connectivity. The consistent negative response of
346 species richness to silt soils may be related to reduced water quality, since sediment nutrient
347 concentrations are higher in finer particle soils (Ockenden et al., 2012), or may reflect the
348 contrasting deposition and landuse histories, with peat soils indicating the historic extent of
349 freshwater marshes and earlier reclamation compared to the marine or riverine deposition of
350 silts that were reclaimed for agriculture more recently. Previous studies have found peat
351 substrates to have distinct flora (Mountford & Arnold, 2006) and support rare invertebrate
352 species (Foster et al., 1989); the richness of species groups in this study were not strongly
353 correlated with peat substrates.

354

355 Environmental factors, such as water quality (Twisk et al., 2000) and flow rate (Leslie et al.,
356 2012), and ditch management type and frequency (Milsom et al., 2004), are known to be
357 important determinants of ditch biodiversity. The inclusion of such variables would certainly
358 improve the predictive power of our models. However, such data were not available at
359 suitable resolution across our study area, and the case is likely to be the same in other
360 regions. We suggest that our predictive modelling approach is used in other regions to predict

361 areas of high potential biodiversity value. Following this, the collation or collection of
362 detailed environmental or habitat data may assist the selection of specific sites for
363 management interventions (such as dredging and cutting), within those areas highlighted by
364 the predictive mapping.

365

366 The effects of many environmental predictors were remarkably consistent among species
367 groups. For example, the richness of all groups was significantly greater closer to existing
368 wetland SSSIs. This may be because the high quality SSSI sites have acted as reservoirs of
369 wetland species, although there may be other conditions not included in this study (e.g. water
370 quality) that are also correlated with the distance to the SSSIs. Whilst the responses of most
371 groups were consistent, the richness of Odonata increased further from the tidal boundary; in
372 contrast, the richness of aquatic species, and wetland and aquatic plants was higher closer to
373 the tidal boundary. This is an important contrast, such that network planning must either take
374 a mixed approach, or select to prioritise either Odonata or remaining groups. Similarly, the
375 value of wooded connectivity networks is rather different for birds, bats and beetles
376 (Boughey et al., 2011; Davies & Pullin, 2007). This adds to previously stated concerns over
377 the use of single taxonomic groups as proxies for other biodiversity (Noss, 1990). Recent
378 work has demonstrated that the addition of habitat characteristics to multi-taxa proxy groups
379 substantially improves the performance of biodiversity surrogates in spatial planning (Di
380 Minin & Moilanen, 2014).

381

382 Biological records can be modelled with environmental variables to predict biodiversity
383 richness across landscapes and such models have been widely used to link species
384 distributions from atlas data to land cover data (e.g. Aauri & de Lucio, 2001; Virkkala et al.,

385 2005). Their use here to model species richness of priority biodiversity across multiple taxa
386 in the Fens allowed previously held assumptions about the importance of landscape features
387 to be tested. However, the use of such methods has been limited by the lack of detailed atlas
388 data for many taxa in many regions, with data for rare and threatened species and for poorly
389 recorded taxonomic groups (i.e. other than vascular plants, butterflies and odonatan)
390 particularly limited. We show that this problem can be overcome by modelling groups of
391 priority taxa with shared ecological requirements, which allowed us to include species that
392 would be too rare and/or sparsely recorded to model individually. This addresses a significant
393 gap in previous large-scale studies that have omitted due to insufficient data, the rare species
394 that are intended to benefit from the conservation measures. Our approach could be applied to
395 any region or landscape where there has been widespread, albeit patchy biological recording.

396

397 We were then able to predict potential species richness, including multi-taxa groups of
398 priority species, at a landscape scale and used the model predictions to make an evidence-
399 based landscape connectivity plan, an improvement on previous plans based on untested
400 expert judgement. Our models predict areas that have the potential for high biodiversity
401 richness, based on their soil and other landscape variables, and we have linked these together
402 with our proposed corridors. However, we do not know if the cells of our predicted corridor
403 currently realise that biodiversity potential with their existing habitat, which could still be
404 improved through enlargement or management, *or* if they currently have low habitat
405 suitability despite high potential on the basis their landscape variables. However, in either
406 case, we predict the potential to enhance biodiversity value and connectivity of those areas to
407 be greater than in areas with lower intrinsic potential and thus we are recommending these
408 areas should be targeted for enhancement.

409

410 The previous attempt to map a strategic connectivity network in the Fens (FFTP 2012)
411 largely concurred with areas of high predicted biodiversity richness. However, our evidence-
412 based map connected a greater proportion of areas with higher potential for biodiversity
413 richness (22% more of our corridors were located in areas of the highest potential richness)
414 and for a shorter overall length. Targeting areas of higher potential richness over a shorter
415 connectivity length is more cost-effective, allowing remaining funding to be targeted to
416 habitat management, a key influence on ditch biodiversity (Milsom et al., 2004). For
417 example, our evidence-based predictive map provides confidence in the strategic targeting of
418 agri-environmental measures and other means to enhance ditch management to those areas of
419 the wider agricultural landscape that have greatest biodiversity potential for aquatic and
420 wetland species.

421

422 Evidence-based predictive models, such as those in this study, could also be further
423 developed to inform optimal connectivity plans. For example, predicted potential species
424 richness can be used as a cost surface for circuit theory and other graph theory based models
425 (Galpern et al., 2011; Rayfield et al., 2011). Although we note that the practical realization of
426 any connectivity plan (subjective or objective) will be dependent on opportunity, landowner
427 and other stakeholder interest, and cost (Bergsten and Zetterberg, 2013), it is crucially
428 important to start negotiations based on evidence. Our methodology utilises ad-hoc records,
429 and thus could be applied in any landscape or region where biological records are available,
430 to provide an evidence-base for network planning, including rare species for which
431 conservation actions are most needed.

References

1. Araujo, M. B., Humphries, C. J., Densham, P. J., Lampinen, R., Hagemeyer, W. J. M., Mitchell-Jones, A. J. & Gasc, J. P. (2001). Would environmental diversity be a good surrogate for species diversity?, *Ecography*, *24*(1), 103-110.
2. Atauri, J. A. & de Lucio, J. V. (2001). The role of landscape structure in species richness distribution of birds, amphibians, reptiles and lepidopterans in Mediterranean landscapes, *Landscape Ecology*, *16*(2), 147-159.
3. Bani, L., Baietto, M., Bottoni, L. & Massa, R. (2002). The use of focal species in designing a habitat network for a lowland area of Lombardy, Italy, *Conservation Biology*, *16*(3), 826-831.
4. Beier, P., Majka, D. R. & Newell, S. L. (2009). Uncertainty analysis of least-cost modeling for designing wildlife linkages, *Ecological Applications*, *19*(8), 2067-2077.
5. Beier, P., Spencer, W., Baldwin, R. F. & McRae, B. H. (2011). Toward best practices for developing regional connectivity maps, *Conservation Biology*, *25*(5), 879-892.
6. Bergsten, A. & Zetterberg, A. (2013). To model the landscape as a network: A practitioner's perspective, *Landscape and Urban Planning*, *119*, 35-43.
7. Boughey, K. L., Lake, I. R., Haysom, K. A. & Dolman, P. M. (2011). Improving the biodiversity benefits of hedgerows: How physical characteristics and the proximity of foraging habitat affect the use of linear features by bats, *Biological Conservation*, *144*(6), 1790-1798.
8. Brooks, T., da Fonseca, G. A. B. & Rodrigues, A. S. L. (2004a). Species, data, and conservation planning, *Conservation Biology*, *18*(6), 1682-1688.
9. Brooks, T. M., da Fonseca, G. A. B., Rodrigues, A. S. L. & (2004b). Protected areas and species, *Conservation Biology*, *18*(3), 616-618.
10. Chen, G., Kéry, M., Plattner, M., Ma, K. & Gardner, B. (2013). Imperfect detection is the rule rather than the exception in plant distribution studies, *Journal of Ecology*, *101*(1), 183-191.
11. Cushman, S. A. & Landguth, E. L. (2012). Multi-taxa population connectivity in the Northern Rocky Mountains, *Ecological Modelling*, *231*, 101-112.
12. Davies, Z. G. & Pullin, A. S. (2007). Are hedgerows effective corridors between fragments of woodland habitat? An evidence-based approach, *Landscape Ecology*, *22*(3), 333-351.
13. Delattre, T., Pichancourt, J. B., Burel, F. & Kindlmann, P. (2010). Grassy field margins as potential corridors for butterflies in agricultural landscapes: A simulation study, *Ecological Modelling*, *221*(2), 370-377.
14. Di Minin, E. & Moilanen, A. (2014). Improving the surrogacy effectiveness of charismatic megafauna with well-surveyed taxonomic groups and habitat types, *Journal of Applied Ecology*, *51*(2), 281-288.

15. Dolman, P. M. (2012). Mechanisms and processes underlying landscape structure effects on bird populations. In R. J. Fuller (Ed), *Birds and Habitat: Relationships in Changing Landscapes* (pp. 93-124), Cambridge University Press, Cambridge, UK.
16. Dolman, P. M., Panter, C. J. & Mossman, H. L. (2012). The biodiversity audit approach challenges regional priorities and identifies a mismatch in conservation, *Journal of Applied Ecology*, 49(5), 986-997.
17. Dorazio, R. M. & Royle, J. A. (2005). Estimating size and composition of biological communities by modeling the occurrence of species, *Journal of the American Statistical Association*, 100(470), 389-398.
18. Eycott, A. E., Marzano, M. & Watts, K. (2011). Filling evidence gaps with expert opinion: The use of Delphi analysis in least-cost modelling of functional connectivity, *Landscape and Urban Planning*, 103(3-4), 400-409.
19. Fens for the Future Partnership (2012) Fens for the Future - a strategic plan for Fenland: a proposal for an enhanced ecological network, Fens for the Future Partnership, Peterborough, UK: <http://www.lincsfenlands.org.uk/index.php?page=BiodiversityFensFuture>.
20. Foster, G. N., Foster, A. P., Eyre, M. D. & Bilton, D. T. (1989). Classification of water beetle assemblages in arable fenland and ranking of sites in relation to conservation value, *Freshwater Biology*, 22(3), 343-354.
21. Freckleton, R. (2002). On the misuse of residuals in ecology: regression of residuals versus multiple regression, *Journal of Animal Ecology*, 71, 42-545.
22. Galpern, P., Manseau, M. & Fall, A. (2011). Patch-based graphs of landscape connectivity: A guide to construction, analysis and application for conservation, *Biological Conservation*, 144(1), 44-55.
23. Herzon, I. & Helenius, J. (2008). Agricultural drainage ditches, their biological importance and functioning, *Biological Conservation*, 141(5), 1171-1183.
24. Hill, J. K., Thomas, C. D., Fox, R., Telfer, M. G., Willis, S. G., Asher, J. & Huntley, B. (2002). Responses of butterflies to twentieth century climate warming: implications for future ranges, *Proceedings of the Royal Society B-Biological Sciences*, 269(1505), 2163-2171.
25. Hill, M. O., Preston, C. D. & Roy, D. B. (2004). PLANTATT-attributes of British and Irish plants: status, size, life history, geography and habitats, Centre for Ecology & Hydrology, Peterborough, UK.
26. Hill, M. O. (2011). Local frequency as a key to interpreting species occurrence data when recording effort is not known, *Methods in Ecology and Evolution*, 3(1), 195-205.
27. Isaac, N. J. B., van Strien, A. J., August, T. A., de Zeeuw, M. P. & Roy, D. B. (2014). Statistics for citizen science: extracting signals of change from noisy ecological data, *Methods in Ecology and Evolution*, 5(10), 1052-1060.
28. Klar, N., Herrmann, M., Henning-Hahn, M., Pott-Doerfer, B., Hofer, H. & Kramer-Schadt, S. (2012). Between ecological theory and planning practice: (Re-) Connecting

- forest patches for the wildcat in Lower Saxony, Germany, *Landscape and Urban Planning*, 105(4), 376-384.
29. Krewenka, K. M., Holzschuh, A., Tschardtke, T. & Dormann, C. F. (2011). Landscape elements as potential barriers and corridors for bees, wasps and parasitoids, *Biological Conservation*, 144(6), 1816-1825.
 30. Kéry, M. (2011). Towards the modelling of true species distributions, *Journal of Biogeography*, 38(4), 617-618.
 31. Lawson, C. R., Bennie, J. J., Thomas, C. D., Hodgson, J. A. & Wilson, R. J. (2012). Local and landscape management of an expanding range margin under climate change, *Journal of Applied Ecology*, 49(3), 552-561.
 32. Lawson, C. R., Bennie, J. J., Thomas, C. D., Hodgson, J. A. & Wilson, R. J. (2014). Active management of protected areas enhances metapopulation expansion under climate change, *Conservation Letters*, 7(2), 111-118.
 33. Legendre, P. & Legendre, L. (2012). Numerical Ecology (3rd edition), In: *Developments in Environmental Modelling*, Elsevier, Amsterdam, The Netherlands.
 34. Leslie, A. W., Smith, R. F., Ruppert, D. E., Bejleri, K., McGrath, J. M., Needelman, B. A. & Lamp, W. O. (2012). Environmental factors structuring benthic macroinvertebrate communities of agricultural ditches in Maryland, *Environmental Entomology*, 41(4), 802-812.
 35. Liira, J. & Paal, T. (2013). Do forest-dwelling plant species disperse along landscape corridors? *Plant Ecology*, 214(3), 455-470.
 36. Lombard, A. T., Cowling, R. M., Pressey, R. L. & Rebelo, A. G. (2003). Effectiveness of land classes as surrogates for species in conservation planning for the Cape Floristic Region, *Biological Conservation*, 112(1-2), 45-62.
 37. MacKenzie, D. I. & Kendall, W. L. (2002). How should detection probability be incorporated into estimates of relative abundance? *Ecology*, 83(9), 2387-2393.
 38. Milsom, T. P., Sherwood, A. J., Rose, S. C., Town, S. J. & Runham, S. R. (2004). Dynamics and management of plant communities in ditches bordering arable fenland in eastern England, *Agriculture Ecosystems & Environment*, 103(1), 85-99.
 39. Morton, R. D., Rowland, C., Wood, C., Meek, L., Marston, C., Smith, G., Wadsworth, R. & Simpson, I. (2011). Land Cover Map 2007 (1km raster dominant Target Class, GB), NERC-Environmental Information Data Centre, Swindon, UK.
 40. Mossman, H. L., Panter, C. J. & Dolman, P. M. (2012). Fens Biodiversity Audit, Report to Fens for the Future steering group, Peterborough, UK.
 41. Mountford, O. & Arnold, H. (2006). Aquatic plant diversity in arable ditches: scoping study, Centre for Ecology and Hydrology, Monks Wood, UK.
 42. Natural England (2013). National Character Area profile: 46. The Fens, Natural England, Peterborough, UK, http://www.naturalengland.org.uk/publications/nca/the_fens.aspx.
 43. Noss, R. F. (1990). Indicators for monitoring biodiversity - a hierarchical approach, *Conservation Biology*, 4(4), 355-364.

44. Nowicki, P., Vrabec, V., Binzenhoefer, B., Feil, J., Zaksek, B., Hovestadt, T. & Settele, J. (2014). Butterfly dispersal in inhospitable matrix: rare, risky, but long-distance, *Landscape Ecology*, 29(3), 401-412.
45. Ockenden, M. C., Deasy, C., Quinton, J. N., Bailey, A. P., Surridge, B. & Stoate, C. (2012). Evaluation of field wetlands for mitigation of diffuse pollution from agriculture: sediment retention, cost and effectiveness, *Environmental Science & Policy*, 24, 110-119.
46. Palmer, M., Drake, M. & Stewart, N. (2013). A manual for the survey and evaluation of the aquatic plant and invertebrate assemblages of grazing marsh ditch systems. Buglife, Peterborough, UK.
47. R Core Team (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
48. Rayfield, B., Fortin, M.-J. & Fall, A. (2011). Connectivity for conservation: a framework to classify network measures, *Ecology*, 92(4), 847-858.
49. Saura, S., Bodin, O. & Fortin, M.-J. (2014). Stepping stones are crucial for species' long-distance dispersal and range expansion through habitat networks, *Journal of Applied Ecology*, 51(1), 171-182.
50. Schindler, S., von Wehrden, H., Poirazidis, K., Wrבka, T. & Kati, V. (2013). Multiscale performance of landscape metrics as indicators of species richness of plants, insects and vertebrates, *Ecological Indicators*, 31, 41-48.
51. Simon, T. N. & Travis, J. (2011). The contribution of man-made ditches to the regional stream biodiversity of the new river watershed in the Florida panhandle, *Hydrobiologia*, 661(1), 163-177.
52. Twisk, W., Noordervliet, M. A. W. & Ter Keurs, W. J. (2000). Effects of ditch management on caddisfly, dragonfly and amphibian larvae in intensively farmed peat areas, *Aquatic Ecology*, 34(4), 397-411.
53. Virkkala, R., Luoto, M., Heikkinen, R. K. & Leikola, N. (2005). Distribution patterns of boreal marshland birds: modelling the relationships to land cover and climate, *Journal of Biogeography*, 32(11), 1957-1970.
54. Williams, P., Whitfield, M., Biggs, J., Bray, S., Fox, G., Nicolet, P. & Sear, D. (2004). Comparative biodiversity of rivers, streams, ditches and ponds in an agricultural landscape in Southern England, *Biological Conservation*, 115(2), 329-341.
55. Zulka, K. P., Abensperg-Traun, M., Milasowszky, N., Bieringer, G., Gereben-Krenn, B.-A., Holzinger, W., Hoelzler, G., Rabitsch, W., Reischuetz, A., Querner, P., Sauberer, N., Schmitzberger, I., Willner, W., Wrבka, T. & Zechmeister, H. (2014). Species richness in dry grassland patches of eastern Austria: A multi-taxon study on the role of local, landscape and habitat quality variables, *Agriculture Ecosystems & Environment*, 182, 25-36.

TABLE LEGENDS

Table 1. Definition and data source of environmental predictors used to model the distribution of Fens biodiversity.

Table 2. Variation in the richness of wetland groups of species explained by the minimum models.

Table 1. Definition and data source of environmental predictors used to model the distribution of Fens biodiversity.

Environmental predictor	Abbreviation used in Figure 2	Source
Mean elevation above sea level: mean elevation of all 50 m x 50 m cells within the 1 km square	Elevation	
Distance to nearest fenland island: Fenland island defined as areas >0.1 km ² with an elevation of ≥ 5 m (excluding coastal cliffs at Skegness and islands within large urban areas). Several large 'islands' within 1000 m of the fenland basin were incorporated into the basin, i.e. not considered islands.	Distance to Fen Island	Edina Digimap Ordnance Survey (OS) PANORAMA DTM (Digital Terrain Model) 1:50,000, 50m cells
Distance to fenland basin: basin was defined as the 5 m contour boundary, unless the area had been defined as a fenland island.		
Presence of either an A or B road within a square	Presence of a road	OS Meridian 2 (1:50 000)
Distance to nearest SSSI comprising wetland habitats	Distance to wetland SSSI	Natural England GIS Digital Boundary Datasets
Percentage of square comprising urban areas. Urban defined from OS Strategic 1:250,000	% urban	
Network distance along 'ditches' to the nearest 'main river'/coastline: calculated using network cost distance. Ditch was defined using the VectorMap District <i>Surface_Water</i> polyline for accurate mapping of small ditches and open water, and the <i>Tidal_Boundary</i> (High/Low Water Mark) polyline because the surface water data stop at the tidal boundary.	Distance to river	Edina Digimap Ordnance Survey Strategic 1:250,000
Network distance along ditch/river to the tidal boundary: calculated using network cost distance (see below for full description). Ditch/river defined using the Edina Digimap <i>River_polyline</i> and VectorMap District <i>Surface_Water</i> polyline. Tidal boundary was defined as the high water mark using the VectorMap District <i>Tidal_Boundary</i> polyline.	Distance to tidal boundary	VectorMap District (1:25,000)
Index length of all ditches per 1 km square: ditches were defined as above. This is considered an index because polylines defined each bank of wide ditches or rivers, resulting in double-counting, as such the lengths are not accurate.	Length of surface water	
Percentage of rough and neutral grassland	% rough/neutral grassland	Land Cover Map 2007. Centre for Ecology and Hydrology
Percentage of grades of Agricultural Land Classification: summed percentage area of grades 3 and 4	% grades 3 & 4	Natural England GIS Digital Boundary Datasets
Percentage of each peat soil type defined using Cranfield Soil Class; Peat; Seasonally wet deep peat to loam; Seasonally wet deep clay over peat (marine alluvium and fen peat) and	% peat; % peaty loam; % deep sand over peat; % deep clay over peat	NATMAP Cranfield University

Seasonally wet deep sand over peat (glaciofluvial drift and peat).

Percentage of silt soil, defined as the Cranfield Soil Class "Seasonally wet deep silty" % silt

Occurrence of a county flora: 0/1 if in a flora recorded county County flora

Table 2. Variation in species richness explained by the minimum models.

	Total r^2	% variation of total r^2 explained		
		Independent effect of recording effort	Independent effect of environment	Joint effect of recording and environment
Odonata	30.4	27.3	52.8	20.0
Fen Specialists	27.2	46.4	27.1	26.4
Aquatic species	30.5	49.2	35.5	15.3
Aquatic plants	59.8	75.4	19.9	4.8
Littoral species	31.5	64.4	33.1	2.5
Wetland species	39.0	57.8	25.5	16.7
Wetland plants	63.9	76.2	17.1	6.7

LIST OF FIGURES

Figure 1. (a) The location of the Fens region within the UK, and b) the intensity of recording effort within the Fens, shown as number of records per 1-km square. Class intervals calculated using Jenks natural breaks.

Figure 2. Mean (\pm SE) standardised effect size (β values) of environmental predictor variables on the richness of ditch indicator groups. Only significant ($p < 0.05$) effects are shown. Predictor abbreviations are provided in Table 1.

Figure 3. Predicted richness per 1-km square for a) Odonata species, b) littoral priority species, c) aquatic priority species, d) aquatic plants, e) wetland priority species, f) wetland plants, and g) Fens Specialists. White areas denote 1 km squares that were excluded from models. Class intervals calculated using Jenks natural breaks.

Figure 4a. Mean of the ranks of predicted species richness per 1 km square of the seven wetland biological indicator groups. A low rank (high number) is given to squares with low predicted biodiversity and high rank (tied, highest = 44) to areas with high biodiversity. Main rivers and (a) connectivity corridors proposed by the Fens for the Future Partnership (excluding the Landscape Corridor) (FFFP 2012), and (b) connectivity corridors re-routed through areas of higher wetland species richness are shown. White areas denote 1 km squares excluded from models. Class intervals calculated using Jenks natural breaks.

Figure 5. Length (km) of primary and secondary connectivity corridors originally proposed by the Fens for the Future Partnership and alternative corridors selectively routed through areas of predicted higher wetland richness. Bars are shaded according to quartiles of the mean of ranks of biodiversity richness per 1-km square across seven indicator groups (Q1: 44-950, Q2: 951-1561, Q3: 1562-2372, Q4: 2373-3000).

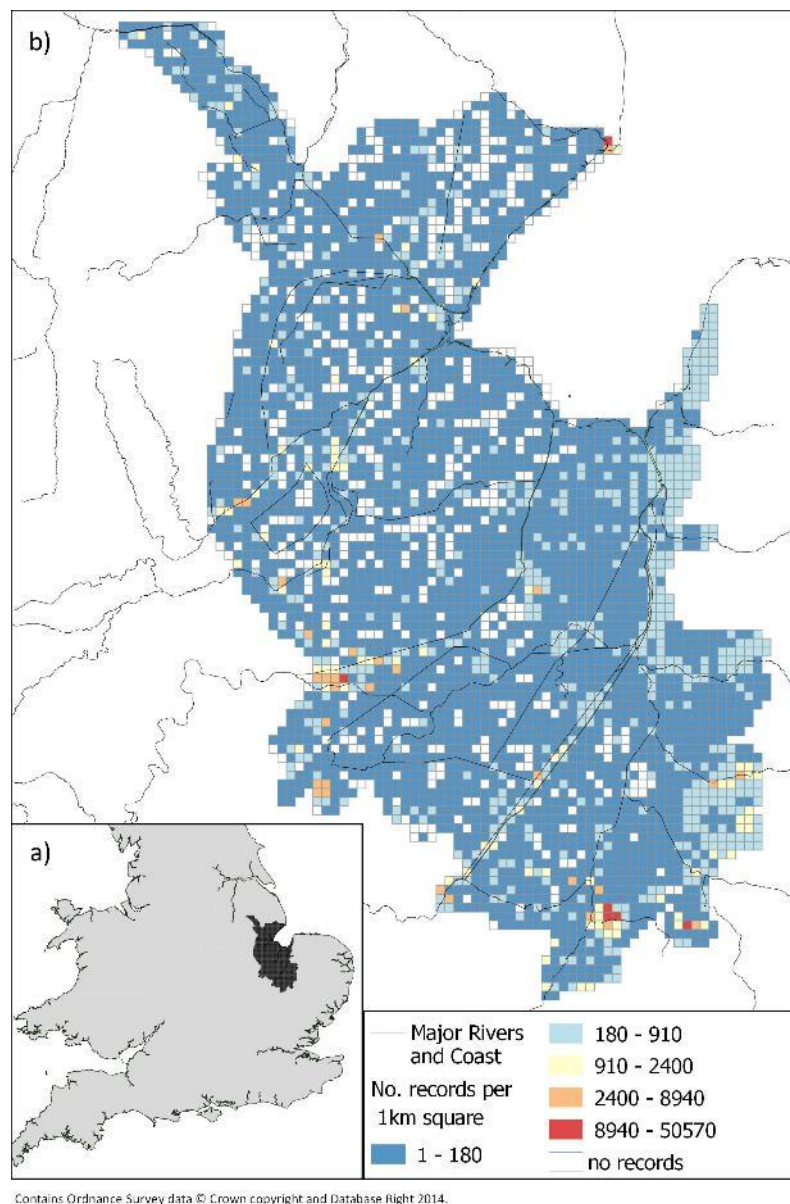


Figure 1. (a) The location of the Fens region within the UK, and b) the intensity of recording effort within the Fens, shown as number of records per 1-km square. Class intervals calculated using Jenks natural breaks.

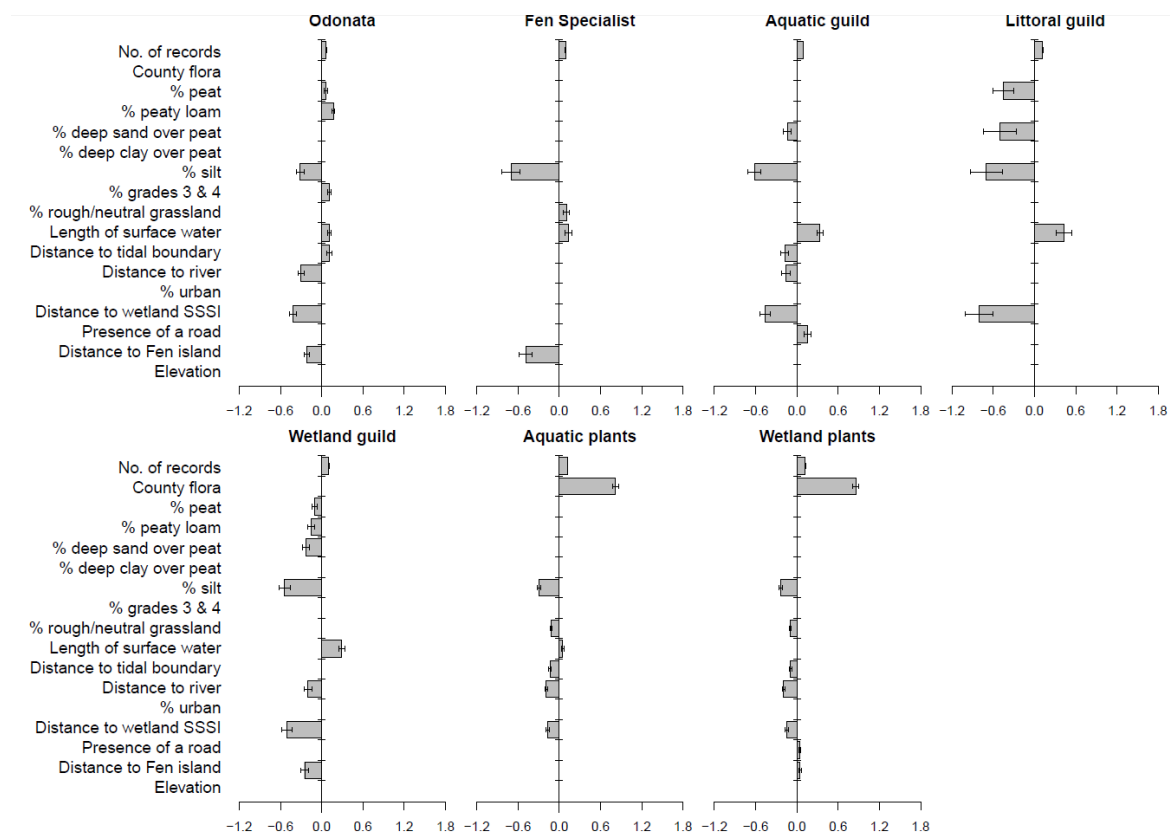


Figure 2. Mean (\pm SE) standardised effect size (β values) of environmental predictor variables on the richness of ditch indicator groups. Only significant ($p < 0.05$) effects are shown. Predictor abbreviations are provided in Table 1.

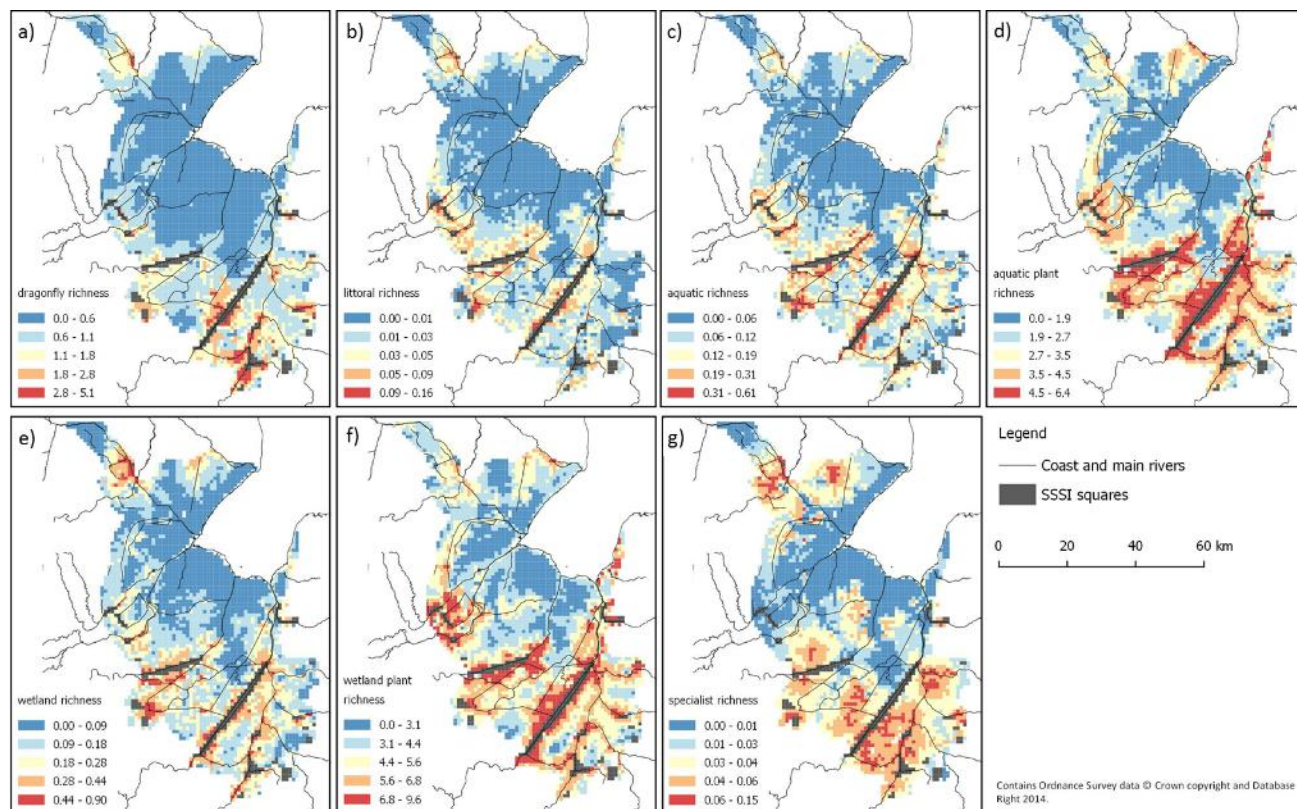
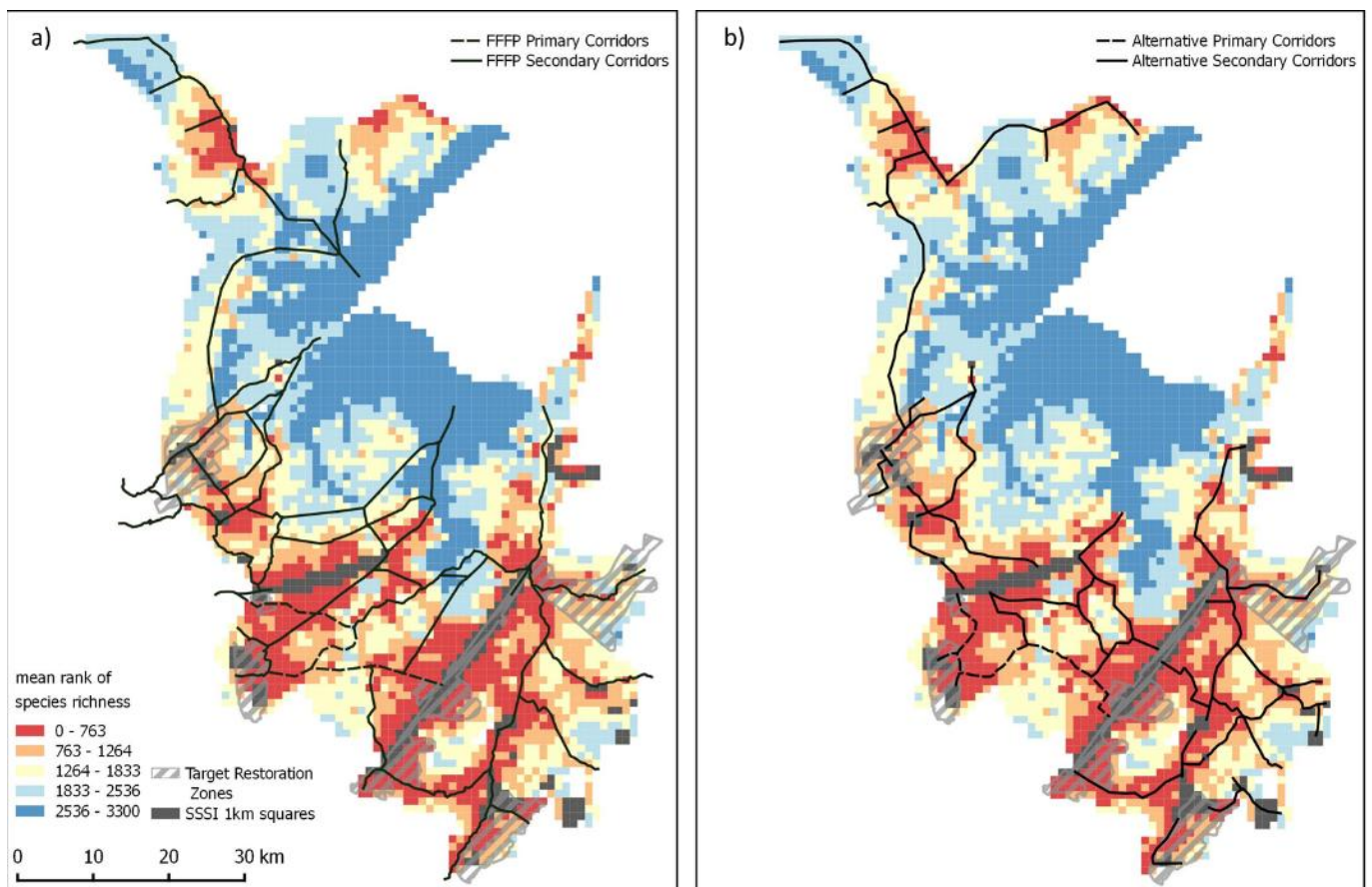


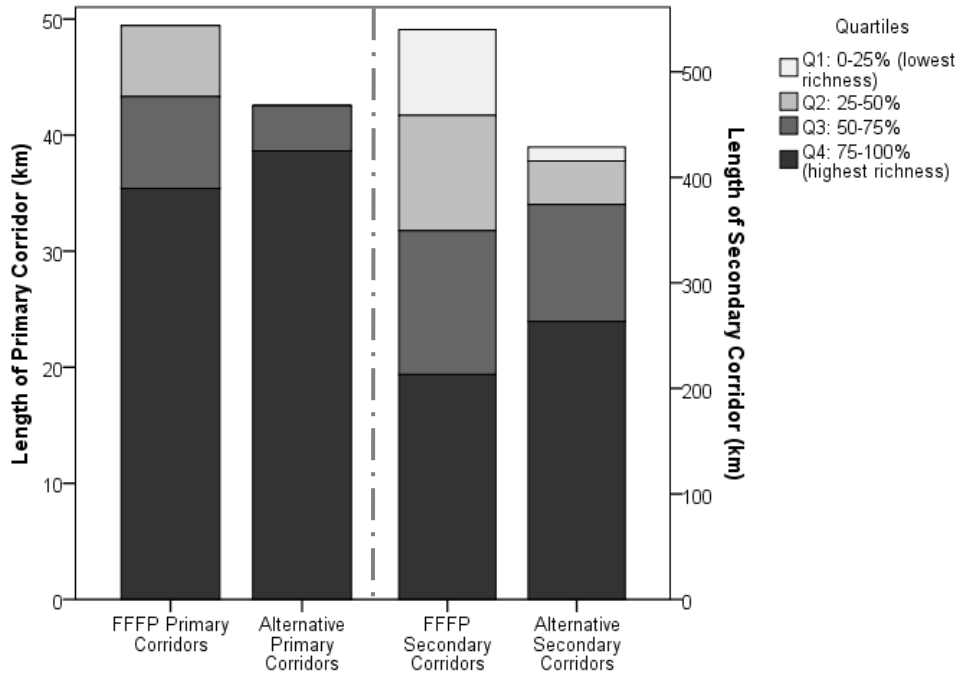
Figure 3. Predicted richness per 1-km square for a) Odonata species, b) littoral priority species, c) aquatic priority species, d) aquatic plants, e) wetland priority species, f) wetland plants, and g) Fens Specialists. White areas denote 1 km squares that were excluded from models. Class intervals calculated using Jenks natural breaks.

Figure 4a. Mean of the ranks of predicted species richness per 1 km square of the seven wetland biological indicator groups. A low rank (high number) is given to squares with low predicted biodiversity and high rank (tied, highest = 44) to areas with high biodiversity. Main rivers and (a) connectivity corridors proposed by the Fens for the Future Partnership (excluding the Landscape Corridor) (FFFP 2012), and (b) connectivity corridors re-routed through areas of higher wetland species richness are shown. White areas denote 1 km squares excluded from models. Class intervals calculated using jenks natural breaks.



Contains Ordnance Survey data © Crown copyright and Database Right 2014.

Figure 5. Length (km) of primary and secondary connectivity corridors originally proposed by the Fens for the Future Partnership and alternative corridors selectively routed through areas of predicted higher wetland richness. Bars are shaded according to quartiles of the mean of ranks of biodiversity richness per 1-km square across seven indicator groups (Q1: 44-950, Q2: 951-1561, Q3: 1562-2372, Q4: 2373-3000).



Acknowledgements

This work is builds on a project funded by the Environment Agency and supported by the Fens for the Future Partnership. We gratefully acknowledge the support and assistance of Natural England, Environment Agency, National Trust, Royal Society for the Protection of Birds, Lincolnshire Wildlife Trust, Bedfordshire, Cambridgeshire and Northamptonshire Wildlife Trust, Cambridge and Peterborough Environmental Records Centre, Suffolk Biological Records Centre, Norfolk Biodiversity Information Service and the Norfolk Biodiversity Partnership. This work would not be possible without the invaluable contributions of hundreds of biological recorders and taxonomic experts. We also acknowledge the statistical advice of Martin Sullivan, and the helpful comments of three anonymous reviewers.