# Magnetic stop signs signal a European songbird's arrival at the breeding site following migration

**Joe Wynn[1*], Oliver Padget[1], Henrik Mouritsen[2,3], Joe Morford[1], Paris Jaggers[1] and Tim Guilford[1*]**

**Affiliations**

1. Oxford Navigation Group, Department of Zoology, 11a Mansfield Road, Oxford, OX1 3SZ, UK

2. AG 'Neurosensorik/Animal Navigation', Carl-von-Ossietzky Universität Oldenburg, 26111 Oldenburg, Germany

3. Research Centre for Neurosensory Sciences, University of Oldenburg, 26111 Oldenburg, Germany

* Correspondence: joseph.wynn@zoo.ox.ac.uk (J.W), tim.guilford@zoo.ox.ac.uk (T.G)

Lead contact**:** joseph.wynn@zoo.ox.ac.uk

**Abstract**

Although it is known that birds can return to their breeding grounds with remarkable precision, it has remained a mystery how they know when and where to stop migrating. Here, using nearly a century of Eurasian reed warbler (*Acrocephalus scirpaceus*) ringing recoveries, we investigated whether fluctuations in the Earth's magnetic field predict variation in the sites to which birds return. Ringing recoveries suggest that magnetic inclination is learned before departure and is subsequently used as a uni-coordinate 'stop sign' when relocating the natal/breeding site. However, many locations have the same inclination angle. Data from populations with different migratory directions indicate that birds solve this ambiguity by stopping at the first place where the right inclination is encountered on an inherited return vector.

**One-sentence summary**

A century of ringing records imply that songbirds use the Earth's magnetic field as a stop sign when relocating their breeding sites.

**Main text**

Whilst first outbound migration in songbirds is typically, though not universally (*1*), thought to utilise an inherited clock and compass (vector) navigation program (*2-4*), additional or alternative mechanisms must be required to account for the accuracy with which birds return to the natal/breeding site ('philopatry'; *5, 6*) . Experienced birds could utilise learnt, bi-coordinate maps comprising 2 or more cue gradients during philopatry, with cues derived from the Earth's magnetic field (EMF) suggested to be prime candidates for accurate natal/breeding site return. (*3, 4*). Such cues include magnetic inclination, the 'dip angle' between the EMF and the Earth's surface; magnetic declination, the angle between magnetic and true north; and magnetic intensity, the overall strength of the EMF (*6*). Maps comprising these components have been suggested to underlie navigation in sea turtles (*7*), fish (*8*), and songbirds (*9, 10*). Previous studies suggest that songbirds could use declination alongside inclination and/or intensity as a bi-coordinate 'map' when compensating for long-distance displacement (*9, 10*), though magnetic gradient cues (alongside other cues) might also allow songbirds to locate the natal/breeding site during philopatry.

Here, we used 17,799 ringing recoveries (1940-2018) made in consecutive years to investigate whether the Eurasian reed warbler (*Acrocephalus scirpaceus*; 'reed warbler'), a trans-Saharan migratory songbird (*11*), utilised magnetic information during philopatry and, if so, how. Because there are slight year-on-year shifts in the EMF, the magnetic parameter values characteristic of an individual's natal/breeding site will exist in a different location the following year. Hence, if birds used magnetic parameters to determine the location of their natal/breeding site, we would expect that positional changes between years reflect year-on-year changes in the location of specific magnetic parameters.

First, we investigated whether changes in specific magnetic parameters (inclination, declination and intensity) between ringing and recovery were smaller than would be expected by chance. This, we reasoned, would test whether birds were attentive to specific magnetic cues. Using randomisation analyses we compared the changes in magnetic field parameters experienced by ringed/recovered reed warblers to those that would be expected if movements were equivalent but random, parameterising our 'null' reed warbler movements using the observed movements (see Figure S3). We found that the median observed change in inclination (randomisation; $p < 0.001$), but not declination ($p > 0.99$) or intensity ($p = 0.18$), was smaller than would be expected under our null model of philopatry. We found, however, that the mean changes in inclination, declination and intensity were not significantly smaller than those expected under our null distribution (randomisation; inclination, $p > 0.99$; declination, $p = 0.4$; intensity, $p = 0.38$). Because the distribution of between-year movements is necessarily long-tailed and non-normal (*12*), it is perhaps unsurprising that the mean and median changes in a given magnetic parameter yield different results. Further, this disparity may imply that the processes driving long-distance dispersal may not be closely linked to magnetic secular variation. To test this, we re-ran our randomisation analysis with birds from the 1$^{st}$ percentile for ringing to recovery distance removed. This was done for adults and chicks separately, ensuring the proportion of adults/chicks was maintained. Using this subset, we found that the observed mean change in inclination was smaller than would be expected by chance (randomisation; $p < 0.001$), but that changes in intensity (randomisation; $p = 0.96$) and declination (randomisation; $p = 0.69$) were not, suggesting our initial result is primarily driven by birds moving shorter distances between ringing and recovery.

Taken together, these results imply that birds on average target the area denoted by the inclination value encountered at the natal/breeding site a year prior. However, since we find no evidence for attention to other magnetic cues, this result is inconsistent with the use of inclination as part of a learnt, bi-coordinate magnetic 'map' during philopatry. Instead, we hypothesise that this is consistent with inclination acting as a uni-coordinate 'stop sign'; birds could recall their natal/breeding location using only one coordinate dimension, if used alongside a compass bearing linking the wintering and breeding sites (*13, 14*).

This would predict that errors in the sites to which birds return should distribute along an inherited return migratory bearing linking the natal/breeding site to the site at which birds crossed the Mediterranean Sea. We reasoned that linking the natal/breeding site to this crossing point was a reasonable estimate for a putative inherited compass bearing, since outbound migration in European passerines is thought to involve a series of migratory bearings, with the most northern bearing thought to connect the breeding site with the Mediterranean crossing point (*11, 15, 16*). Reed warblers breeding in northern/western Europe cross the Mediterranean in southwestern Europe, whilst birds breeding southeast of the Austrian Alps are known to cross the Mediterranean in southeast Europe (*11, 17, 18*). Therefore, we predict that if reed warblers utilise magnetic stop signs during philopatry, changes in longitude/latitude between consecutive breeding seasons should distribute along a southwest-to-northeast axis for northern/western birds and along a southeast-to-northwest axis for southeastern birds.

When changes in latitude were regressed against changes in longitude, we found a significant change in latitude with longitude (LM; $F = 143{,}445$, $p < 0.0001$) and, further, found a significant

difference in the direction of this effect between migratory populations (LM; $F = 1,579$, $p < 0.0001$; see Figure 1). We found that the northern/western populations showed significant eastwards movement with northern movement (gradient = $1.51 \pm 0.0480$), whilst the southeastern population showed significant westwards movement with northwards movement (gradient = $-1.21 \pm 0.0675$; see Figure 1). This is consistent with birds using magnetic inclination as a stop sign, though is also consistent with ineffective exclusion of non-breeding birds still currently migrating. Therefore, we sought to ensure the result was robust to our selection criteria by re-running our analyses with 3 more stringent subsets: birds ringed in a restricted core breeding window (15th June-15th July); birds ringed as chicks and recovered at/north of their natal site; and birds sexed at capture, which requires identification via breeding phenotype (*19*; see Figures S4-S7).

Under a magnetic stop sign hypothesis, we might expect birds to be recovered at the intersect between the natal/breeding inclination isoline and the return compass bearing. A similarly precise predicted recovery location exists for any magnetic navigation hypothesis, including both stop sign and bi-coordinate models (see Figure 2). Since for each bird there is a unique location where it would be expected to be recovered under a given model of navigation, we can measure how close a given ringing recovery is to this position (the 'observed-versus-expected distance'; see Figure 2; see Methods). The smaller this observed-versus-expected distance is, the better the hypothesis fits the observed data. We can test the likelihood of different models of navigation in a randomisation by comparing observed-versus-expected distances for a given hypothesis to the observed-versus-expected distances of 'null birds' making equivalent but random between-year movements (see Figure S9).

We found that the only hypothesis that significantly out-performed the null model was inclination acting as a stop-sign (randomisation; $p < 0.001$; see Figures S1/S2), with the observed-versus-expected distance of all other models no smaller than would be expected if birds moved randomly. Remarkably, we also found that birds were recovered closer to the site predicted by the inclination stop-sign model than they were to their natal/breeding site (Mann-Whitney U test; $p < 0.0001$), implying that birds showed preference for the site predicted by the inclination stop-sign over even their own breeding/natal site.

Since this is a correlative analysis it is, necessarily, possible that our findings are the result of a confound between movements of the magnetic field and some other parameter. However, given that we found no substantial confound with environmental phenology (see Table S1), we found no difference in our results when analysing different subsets of the data (see Supplementary Text) and there are no long-term temporal trends in longitudinal/latitudinal shift between ringing and recovery (latitude LM; $F = 1.69$, $p = 0.190$; longitude LM; $F = 0.99$, $p = 0.590$), we believe that the most parsimonious interpretation of our result is that magnetic inclination is used as a stop sign during philopatry. This might make sense, since magnetic inclination has been repeatedly implicated in avian navigation (*6*). Additionally, other magnetic gradient-derived positions move further with secular variation, thus making the proposed mechanism relatively robust. The position of the natal site as estimated using inclination and declination as a bi-coordinate map would move, on average, 18.5km (± 0.0760km) between years, intensity and declination 20.4km (± 0.0510km) and intensity and inclination 98.2km (± 2.60km). In contrast, the location of the breeding site denoted using inclination as a stop sign moves only 1.22km (± 0.0133km) between years. We suggest that, by remembering breeding location relative to the most stable cue, and referencing it

alongside a compass bearing, the proposed strategy minimises the impact of secular variation. Nevertheless, an inclination stop sign cannot be the only system used, since even slight secular variation would inhibit philopatry. Therefore, other cues must complement magnetic inclination when locating the natal/breeding site (*6*).

Whilst our results shed light on the sensory and developmental underpinnings of philopatry, they may also imply that magnetic secular variation is of some importance when considering the drivers of range shift in migratory taxa (*20, 21*). As with any correlative contrasts drawn using environmental variance, experimental verification of our suggested mechanism is essential. Nonetheless, we believe our findings provide evidence for a novel mechanism of long-distance navigation, both within birds and migratory animals more generally.

**Acknowledgements**

We would like to thank everyone across Europe who has contributed to this fantastic dataset over the years: this study would not be possible without their input. We also thank the members of the Oxnav research group for their insightful contributions, and we thank Amy Holguin, Chlöe Randall and Tom Malpas for their comments on a draft manuscript.

**Funding**

J.W. and J.M. were funded by a UKRI BBSRC scholarship [grant number BB/M011224/1], P.J. was funded by a UKRI NERC scholarship [grant number NE/S007474/1], O.P. was funded by a Junior Research Fellowship at St John's College, University of Oxford, H.M. received funding from the European Research Council (under the European Union's Horizon 2020 research and innovation programme, grant agreement no. 810002 (Synergy Grant: "QuantumBirds")) and the Deutsche Forschungsgemeinschaft (SFB 1372, "Magnetoreception and navigation in vertebrates" and GRK 1885, "Molecular basis of sensory biology") and T.G.'s research was supported by Merton College, Oxford, and by the Mary Griffiths award.

## Author Contributions

Conceptualization: JW, OP, TG

Methodology: JW, OP

Investigation: JW, OP, TG, HM, JM, PJ

Writing – original draft: JW

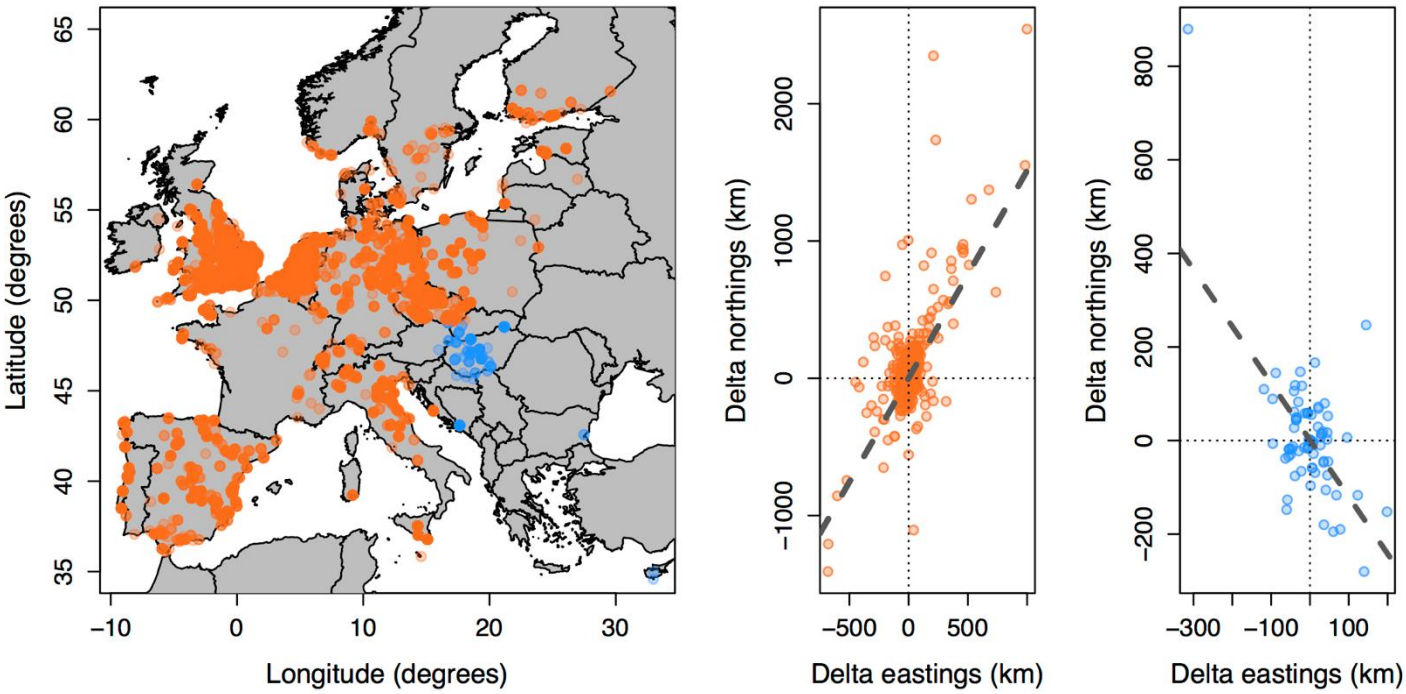
Writing – review & editing: JW, OP, TG, HM, JM, PJ

## Competing interests

The authors declare no competing interests.

## Data and materials availability

Data are available from the Dryad repository <REPOSITORY ID>.




`

**Figure 1: Changes in position between ringing and recovery for Eurasian reed warblers.** (left) Reed warbler ringing sites, coloured orange for west-migrating birds and blue for east-migrating birds. (middle) Changes in longitude ('eastings' in kilometres) plotted against changes in latitude ('northings' in kilometres) for birds breeding in northern/western Europe and (right) changes in longitude plotted against changes in latitude for birds breeding in southeastern Europe. In all panels, opacity indicates multiple overlapping datapoints.

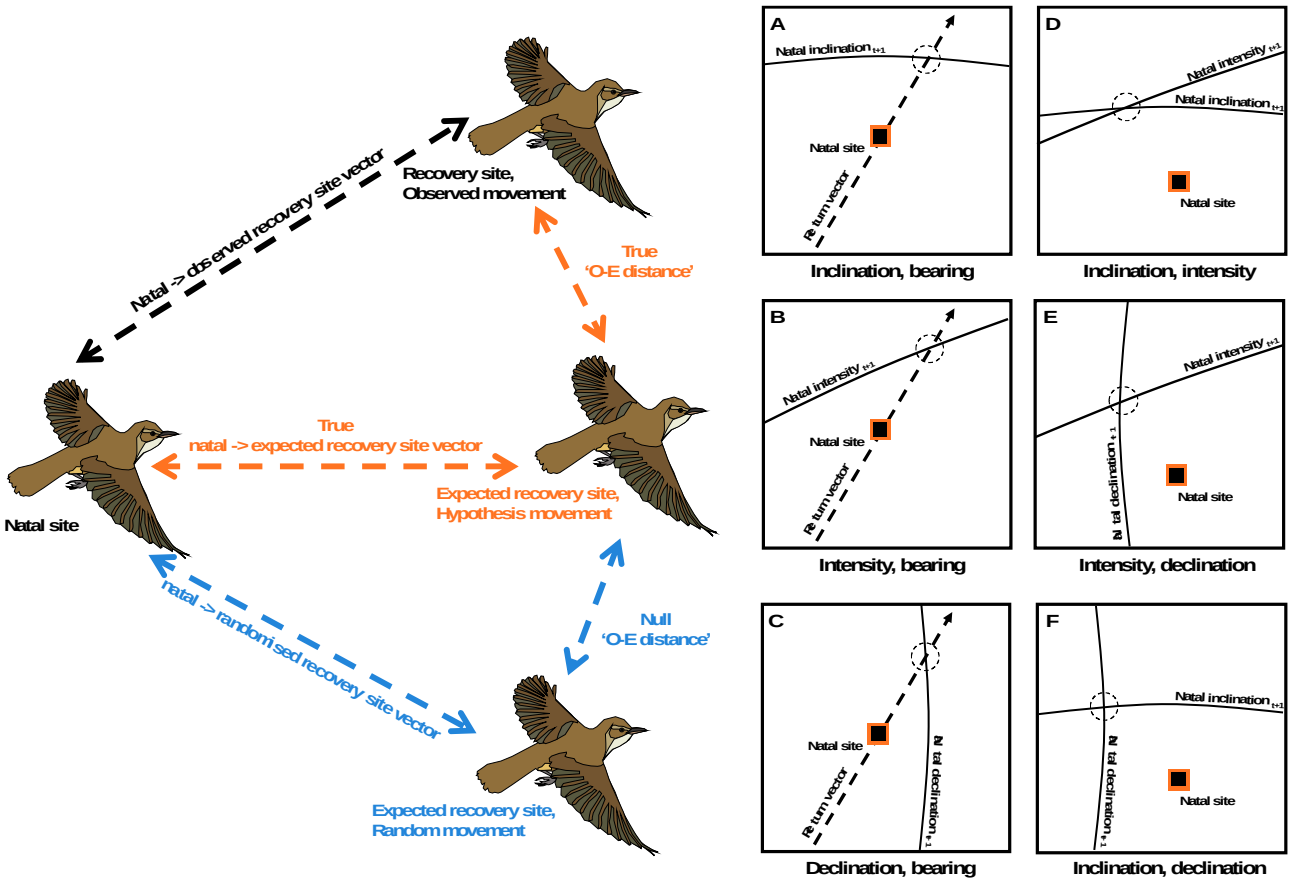
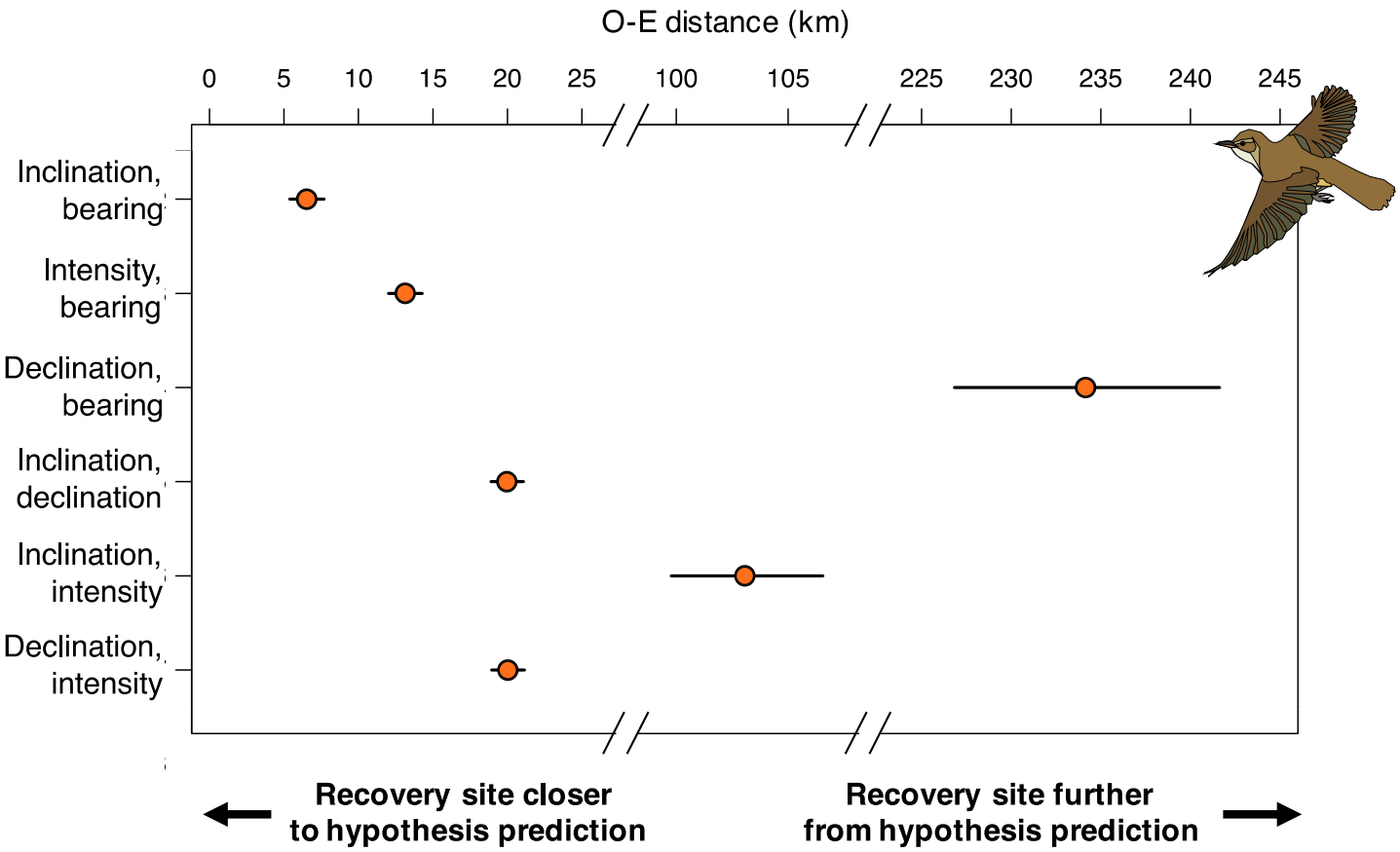



**Figure 2: Using the observed-versus-expected distance to test navigational hypotheses.** (left) The observed-versus-expected distance between the hypothesised and observed recovery sites ('True O-E distance') can be calculated, and compared to the distance between the hypothesised recovery site and the recovery site expected under equal but random movements ('null O-E distance'). Through this comparison we can calculate the likelihood of a hypothesis out-performing chance. (right) Panels A-F show how expected recovery positions under each hypothesis are calculated.




206 **Figure 3: How closely do different navigational hypotheses fit the observed ringing data?** A
207 comparison of the mean observed-versus-expected distances (see Figure 2) for all hypotheses
208 tested, with smaller values representing a better fit between the hypothesis and the observed data.
209 Error bars represent bootstrapped 99% confidence intervals, note the breaks in the scale on the
210 horizontal axis.

211

# Supplementary materials

## Materials and methods

### Ringing data

Reed warblers are night-migrating passerine songbirds breeding in wetland habitats across the temperate Western Palearctic from Iberia to western Russia (*11*). Ringing data were derived from the EURING scheme (euring.org) based on a search query for any reed warblers ringed and recovered in subsequent years. Many reed warblers are ringed not at the breeding site but instead on migration. As such, ringing/recovery events were filtered so as to restrict the records used to breeding birds. Previous studies suggest that reed warblers commence breeding in late May in southern Europe and mid-June in northern Europe (*22, 23*) whilst ringing totals at coastal sites both in northern Europe (Portland Bird Observatory; Southern England) and southern Europe (Gibraltar Bird Observatory; Iberia) do not record migrant reed warblers in large numbers until August. Consequently, we believe a late July cut-off is sufficient to remove the majority of migrant birds. However, it is possible that birds ringed/recovered in May and June are on northward migration. To remove these birds, we used a linear regression model of Julian date against longitude and latitude trained on ringing records of adults with a brood patch in order to estimate the point at which breeding likely commenced at a given longitude/latitude. Based on this model we estimated, for each ringing/recovery's longitude/latitude pairing, a cut-off Julian date where breeding was likely to commence at that specific location. We then removed ringing records that fell before this date, reasoning that such records may represent birds on migration rather than breeding birds. For the birds used in our study, cut-off dates varied between the 16$^{th}$ June (in the southernmost latitudes of the range) and 8$^{th}$ July (in the northernmost latitudes). Through this process, some 75,000+ ringing recoveries were reduced to 17,799 where the ringing event occurred

after the cut-off date (but before 31st July). This subset broke down into 10,784 birds ringed and recovered as adults, and 7,015 birds ringed as nestlings/fledglings and recovered as second calendar year adults. In total, our analysis used birds ringed between the years 1940 and 2018.

The above outlined latitude/longitude-specific cut-off date allows us to select for birds likely to be breeding, whilst simultaneously allowing for the largest possible sample size. Consequently, we believe that this mechanism represents the optimum trade-off between maximising sample size and minimising the probability of erroneously including migrant birds in our dataset. However, ensuring that any results are robust to the precise selection process used is essential. We have, therefore, conducted re-analyses of 3 further subsets of our data selected using different criteria. These subsets are birds captured/recovered between 15th June - 15th July, birds for which a sex had been determined (at both the ringing and recovery events) and nestlings subsequently recovered as adults at or north of the breeding site. For the first of our subsets we reasoned that it was extremely unlikely these birds were on migration given previous studies into reed warbler breeding phenology (*22, 23*). For our second subset we reasoned that, given sex-determination in reed warblers requires examination of breeding-specific phenotypical features (brood patch or cloacal protrusion; *19*), sexed birds were very likely to be recovered at the breeding site. We used positive sex identification, rather than the presence/absence of a brood patch/cloacal protrusion, as our criteria for inclusion in this analysis because the latter is not widely reported in the EURING data. For our fourth subset, we reasoned that for nestlings the ringing event had to take place at the breeding site (by definition), and that birds captured at or north of the breeding site were extremely unlikely to be still be on northwards migration when re-encountered. Therefore, by analysing these additional subsets in addition to our main dataset, we sought to ensure that the results presented in

our main text were robust to the selection process used to determine which ringing recoveries to use and, further, ensure that our findings did not result from the inclusion of migrating birds in the analysis dataset.

In addition to testing subsets of our data selected using different selection criteria, we also reanalysed our data including a 'constant effort' cohort so as to ensure our result was not driven by a handful of year's data and was instead caused by a trend observed across all years. Our constant effort cohort comprised 50 ringing records selected at random from every year since 1967, the first year for which we have 50 records, upon which we repeated all our analyses. These are presented in the supplementary material.

A not insubstantial number of reed warblers (n = 1,321) have been caught more than twice in consecutive years, and hence there is the possibility that the inclusion of such records in our analysis might be pseudoreplicative. To investigate whether this were the case, prior to our main analysis we regressed the distance between the first encounter site and the second encounter site against the distance between the second encounter site and the third encounter site. These distances were used as a proxy for philopatric faithfulness and, if faithfulness was consistent between years, we reasoned multiple records from the same bird could not be included in our analysis. We found, however, that this was not the case (LM; $F = 0.0462$, $p = 0.830$). Consequently, since we found no within-individual consistency in movements between years, we reasoned that philopatry in one year was unlikely to predict philopatry in the next and hence all records could be included in our analysis.

**Magnetic and environmental data**

Magnetic data were derived from the International Geomagnetic Reference Field 12 at a spatial resolution of 0.2º longitude and 0.2º latitude. Bi-linear interpolation was used to resample magnetic parameters onto a precise 2.5km x 2.5km grid, so as the grid resolution was constant throughout the range. For each year included in the study, magnetic parameters were sampled on a 2-weekly basis, and from the 16 values obtained per year a mean was derived at each location for each of inclination, declination and intensity. Using this, each ringing/recovery event was assigned a magnetic inclination, intensity and declination value based on the closest possible grid cell (maximum distance to nearest grid cell = 1.25km).

In addition, a total of 4 environmental variables were derived and their correlation against different magnetic variables investigated. These were: surface air temperature, the normalized difference vegetation index (NDVI, a proxy for live vegetative cover), the Fraction of Absorbed Photosynthetically Active Radiation (fAPAR, a proxy for live vegetative cover) and the proportion of days where a temperature of < 0ºC was recorded ('proportion frost days'). All environmental variables were gathered over the largest possible temporal range for all positions across Europe,and stored as a raster grid. For each grid cell, data were gathered for dates in May-July and a mean value for each position was calculated to represent the environmental phenology in that area in a given year. Temperature data were derived from the National Center for Environmental Prediction reanalysis dataset at a spatial resolution of 0.5º longitude and 0.5º latitude. NDVI was derived from the NASA AppEEARS portal for the years 2000-2020 at a spatial resolution of 1km x 1km. The fAPAR and the proportion of frost days were derived from the European Copernicus reanalysis, and were sampled at a spatial resolution of 1km x 1km over the years 1981-2014.

Each magnetic/environmental variable was normalised and regressed against all other magnetic and environmental variables in a series of uni-variate linear regressions. Because each variable was normalised, a gradient of +1 indicated a strong positive correlation, a gradient of -1 indicated a strong negative correlation and a gradient of 0 indicated no correlation. These gradients were presented in a matrix, and used to investigate the likelihood that any significant correlations between philopatric movements and magnetic variables were driven by a confound between magnetic and environmental variables.

**Statistics and analysis**

All statistics were conducted in R (R *24*).

*Parameter shift randomisation analysis*

We might expect that if birds attended to specific magnetic cues, they might actively minimise the change in specific magnetic parameters between ringing and recovery. We would, therefore, expect that the observed magnetic parameter shift between ringing and recovery would be smaller in the focal bird's observed movement than in equivalent but random movements. As such, we reasoned that we can test whether birds actively attend to magnetic parameters by comparing their movements to realistic but random null 'movements' in a randomisation analysis.

It is essential, therefore, that a realistic 'null' expectation is established for reed warbler philopatry. Based on previous studies, we may expect birds to be recovered close to their natal/breeding site (*25*). However, the potential value of appropriate habitat (rather than precise location) when

considering fitness maximisation, coupled to any potential inbreeding costs, the theoretical optimal precision of philopatry is unclear (*5*). Consequently, we derived our null model for natal/breeding philopatry from the empirically gathered ringing recoveries, so as the philopatry of our null birds reflected that of real birds.

Additionally, not all changes in position are available to all birds and it is essential that any null model takes this into account. For example, if a reed warbler attempting to return to the Netherlands encountered a northward shift in the Earth's magnetic field then its expected recovery position, under a model of solely magnetic navigation, would likely end up in the North Sea. This is impossible as, even if birds did respond precisely to such magnetic shifts, they would never be recovered owing to a lack of offshore ringing effort. Hence, we might expect that, given the topography of the coastline and the distribution of ringing effort, the most likely recovery position is on the nearest adjacent land. This might result in unpredictable changes in distribution with magnetic parameters; to take the above example, the longitude may change with inclination alongside latitude due to the shape of the coastline. It is, therefore, essential that the null distribution against which any hypotheses are tested takes into account not only the topography of Europe but, additionally, the patchy distribution of ringing effort over the reed warbler breeding range. This can be achieved by limiting 'null' recoveries to the sites at which real birds have been encountered.

To establish null expected recovery positions for a given bird we first sampled, with replacement, the empirically observed distances between ringing and recovery, so as our null reed warblers moved the same distances as real reed warblers. The distances used to generate null positions in

adult birds were only drawn from distances derived from other adult birds, and likewise the distances used to generate null positions in 1st year birds were only drawn from distances derived from other 1st years. This meant that, if there were differences in the accuracy of homing between adult and 1st year reed warblers, these were taken into account in our null distribution. For each bird we then selected a recovery site from the sites at which reed warblers had been ringed/recovered that were within 1km of the sampled between-year distance. If no such point existed, we selected the point that was closest in distance to the sampled between-year distance. This ensured that every null recovery site was an observed recovery location, hence our null model didn't include unrealistic recovery sites (such as over the sea or in an area with no ringing effort; see Figure S9). In turn, this allowed us to account for the patchy distribution of ringing effort across Europe. Since the longitude and latitude at which ringing occurred didn't change significantly over time (latitude LM; $F = 1.69$, $p = 0.190$; longitude LM; $F = 0.99$, $p = 0.590$), and because changes in the longitude/latitude at which birds were ringed/recovered did not predict changes in longitude/latitude between ringing and recovery (latitude LM; $F = 1.59$, $p = 0.201$; longitude LM; $F = 0.672$, $p = 0.413$), we allowed our null birds to be recovered at any site reed warblers had been encountered at irrespective of the dates at which the site was active. This allowed for the greatest possible variation in the movement of our null birds.

For each null 'movement' we calculated the associated change in magnetic parameters. The mean and median changes in the parameters of interest were then calculated for all birds per iteration. This process was repeated 1,000 times, and the distribution of median/mean changes in the magnetic field per iteration were used as our null distribution. If birds were using the parameter in question to guide homewards movement, we reasoned that the true median and/or mean change in

magnetic parameters derived from real ringing/recovery events would be significantly smaller than those generated by simulating random but equivalent movements. Thus, by assessing the proportion of times the null median/mean shift in a given magnetic parameter was smaller than the observed median/mean shift, we could ascertain whether a statistically significant preference for magnetic inclination, declination or intensity was observed.

*Observed-versus-expected comparison analysis*

Under any hypothesis of magnetic navigation defined a priori, there is a precise location at which a bird is expected to be observed. For example, a bird navigating using a bi-coordinate map comprising magnetic inclination and magnetic declination would be expected to be recovered at the intersect of the natal inclination/declination isolines observed in the subsequent year. As such, we can test how birds use specific cues by comparing their recovery location to that expected under a given hypothesis.

To do this, we attempted to test specifically whether the movement of reed warblers between ringing and recovery correlated with 'movement' in the site occupied by the magnetic parameter values previously found at the natal/breeding site in the previous year. For instance, if the latitude and longitude of magnetic values $X_1$ $Y_1$ (e.g. inclination/declination) moved between years, there would be some correlation between this movement and the movement between ringing and recovery of birds if they aimed specifically for coordinates in the cues hypothesised. To test whether this was the case, we isolated the exact geographic position to which a bird is expected to home under a specific model of navigation. In our analysis we included 6 potential magnetic

hypotheses, each containing either 2 magnetic parameters used as a bi-coordinate map or a single magnetic parameter utilised alongside a return compass bearing as a stop sign. These hypotheses were: inclination/declination, inclination/intensity, intensity/declination, inclination/bearing ('inclination stop sign'), declination/bearing ('declination stop sign') and intensity/bearing ('intensity stop sign').

For bi-coordinate magnetic map hypotheses we defined the expected recovery position as the intersect between the isolines denoting the magnetic parameter values associated with the natal/breeding site in the natal/breeding year. For the stop sign hypotheses a likely compass bearing was derived for each bird by connecting each bird's natal/breeding site with the likely point at which each bird crossed the Mediterranean Sea (*11, 18*), with the resultant line's intersect with the magnetic isoline of interest giving the predicted recovery location for each bird. We reasoned that linking the natal/breeding site to the point at which each bird was likely to cross the Mediterranean Sea was a reasonable estimate for a putative compass bearing, since birds on outbound migration are thought to navigate along a series of inherited migratory bearings, the most northern of which is thought to connect the natal/breeding site with the Mediterranean crossing point (*15, 16*). If a bird were to re-capitulate a series of innate bearings we would, therefore, expect the final leg of such a series would intersect the Mediterranean Sea crossing site. For birds breeding in western and northern Europe the Mediterranean crossing point was estimated as the Straits of Gibraltar, and for birds breeding in south-eastern Europe the Mediterranean crossing point was estimated as the western coast of Lebanon. Whilst it is extremely unlikely that all birds pass precisely through these positions we reasoned that, given how great the distance is between most breeding sites and

the crossing points, even relatively large discrepancies between the hypothesised and observed crossing points will have a fairly small impact on any subsequent analysis.

If, for a given hypothesis, the expected recovery site was at a probabilistically impossible site (for example in a mountain range, over the sea or in an area where no ringing occurred) we reasoned it would be unrealistic to suggest that such a site is the 'expected' recovery site. As such, for each hypothesis we defined the expected recovery site as the site where ringing occurred closest to the site where a bird was expected to be recovered under a given hypothesis.

For each bird we measured the geographic distance between the observed recovery site and the recovery site expected under a given model of navigation (the 'observed-versus-expected'distance), a process which we repeated for each navigational hypothesis. The smaller this distance, the better the fit between the ringing data and the hypothesis in question. In order to assess the relative likelihood of different magnetic models of navigation, we compared the observed-versus-expected distance for each bird to the distances between the expected recovery site and a null recovery site (the 'null observed-versus-expected distance'; see above for a description of our null model). To test each hypothesis against the null distribution, we iteratively generated a null recovery site for each bird using the above protocol (n = 1000). For each iteration we calculated a mean null observed-versus-expected distance for each hypothesis, and the distribution of null observed-versus-expected distances were used to test whether a given hypothesis' real observed-versus-expected distance was smaller than would be expected under the null distributions (i.e. was smaller than would be expected if birds were inattentive to the Earth's

magnetic field). P-values were calculated as the proportion of null mean observed-versus-expected distances that were smaller than the observed mean observed-versus-expected distance.

**Supplementary text**

*Age-specific observed-versus-expected analysis*

Despite 1st years moving further between ringing and recovery (mean = 11.51km ± 0.607) than adults (mean = 4.388km ± 0.390), consistent with previous studies, we found that the observed-versus-expected distance for the inclination stop sign hypothesis was significantly smaller than chance when considering both adults (randomisation; $p < 0.001$) and 1st years (randomisation; $p < 0.001$) separately, implying that even though adult birds were more accurate in their homing birds of all ages attended to magnetic stop signs.

*Multiple subset analyses*

In addition to our main analyses, we repeated all statistical tests using different subsets of the ringing data, so as to ensure that the trends discussed in our main text were robust to the selection criteria used to subset ringing data.

In total, we conducted 4 additional analyses. Firstly, we repeated our analyses for a subset of birds ringed at the nest and recovered at or north of the breeding site. Such birds, we reasoned, were extremely unlikely to represent birds caught on migration in either instance. Based on this subset we found, consistent with our main analyses, that the shift in median inclination (randomisation; $p < 0.001$; see methods) but not declination (randomisation; $p = 0.87$) or intensity (randomisation; $p = 0.13$; see figure S4A) was smaller than would be expected if birds had moved randomly. We additionally found a significant change in latitude per unit longitude (LM; $F = 4,360$, $p < 0.0001$) and, again consistent with our main analysis, we found that the direction of this effect was different for different breeding populations (south-eastern vs north/western Europe; LM; $F = 47.4$, $p <$

466 0.0001; see figure S4B). Finally, as in our main analysis, we found that birds were significantly
467 closer to the expected recovery location under an inclination/bearing model of navigation than
468 would be expected by chance (randomisation; $p < 0.001$; see methods; see figure S4C). This was
469 not the case for any other hypothesis examined (see figure S4C).
470
471 Our second subset analysis focussed on birds caught between the 15$^{th}$ June and 15$^{th}$ July. Based
472 on previous studies, this date range represents the most likely breeding window for Eurasian reed
473 warblers. Therefore, this is when birds are least likely to be on migration (*22*, *23*). When
474 considering birds ringed between these dates, consistent with our main analyses, we found that the
475 shift in median inclination (randomisation; $p < 0.001$; see methods) was smaller than would be
476 expected if birds had moved randomly. This was not the case when considering declination
477 (randomisation; $p = 0.87$) or intensity (randomisation $p = 0.13$; see figure S5A). Further, we found
478 that changes in longitude were significantly correlated with changes in latitude (LM; $F = 7{,}576$ , $p$
479 $< 0.0001$) and, again consistent with our main analysis, we found that the direction of this effect
480 was different between the south-eastern and north/western European reed warbler populations
481 (LM; $F = 554$, $p < 0.0001$; see figure S5B). Finally, we found that birds were significantly closer
482 to the expected recovery location under an inclination/bearing model of navigation than would be
483 expected by chance (randomisation; $p < 0.001$; see methods; see figure S5C). This was not the
484 case for any other hypothesis examined (see figure S5C).
485
486 Our third subset analysis concerned birds for which sex had been determined. Given the overlap
487 in plumage and biometrics between male and female reed warblers, sexing typically relies on the
488 presence of a brood patch/cloacal protrusion (*19*). Sexed birds are, therefore, likely to be breeding.

Amongst birds sexed at both ringing and recovery, we found that the shift in median inclination (randomisation; $p < 0.001$; see methods) was smaller than would be expected if birds had moved randomly. This was not the case when considering declination (randomisation $p = 0.99$) or intensity (randomisation $p = 0.47$; see figure S6A). Additionally, we found that changes in longitude were significantly predicted by changes in latitude (LM; $F = 12{,}196$, $p < 0.0001$) and, again consistent with our main analysis, we found that the direction of this effect was different between the south-eastern and north/western European reed warbler populations (LM; $F = 34{,}349$, $p < 0.0001$; see figure S6B). Finally, we found that birds were significantly closer to the expected recovery location under an inclination/bearing model of navigation than would be expected by chance (randomisation; $p = 0.006$; see methods; see figure S6C). This was not the case for any other hypothesis examined (see figure S6C).

Additionally, we further repeated our analysis of sexed birds that were also included in our original analysis (i.e. were both sexed and additionally ringed/recovered after a cut-off date specified for their ringing longitude/latitude). Amongst these birds we again found that the shift in median inclination (randomisation; $p < 0.001$; see methods) was smaller than would be expected if birds had moved randomly. This was not the case when considering declination (randomisation; $p > 0.99$) or intensity (randomisation $p = 0.48$; see figure S7A). Further, we found that changes in longitude were significantly correlated with changes in latitude (LM; $F = 46.7$, $p < 0.0001$) and, again consistent with our main analysis, we found that the direction of this effect was different between the south-eastern and north/western European reed warbler populations (LM; $F = 26.9$, $p < 0.0001$; see figure S7B). Finally, we found that birds were significantly closer to the expected recovery location under an inclination/bearing model of navigation than would be expected by

chance (randomisation; $p < 0.001$; see methods; see figure S6C). This was not the case for any other hypothesis examined (see figure S7C).

Our final subset analysis concerned a 'constant effort' cohort, including 50 birds ringed every year since 1965 (the first year in which 50 birds were ringed and subsequently recovered). This analysis was included so as to ensure our results were not disproportionately influenced by a handful of years' data. This is particularly important given that crustal anomalies, diurnal variation, or variability in the magnetic field owing to space weather are not modeled completely by the IGRF. Amongst these birds we again found that the shift in median inclination (randomisation; $p < 0.001$; see methods) was smaller than would be expected if birds had moved randomly. This was not the case when considering declination (randomisation; $p = 0.94$) or intensity (randomisation $p = 0.010$; see figure S8A). Further, we found that changes in longitude were significantly correlated with changes in latitude (LM; $F = 46.7$, $p < 0.0001$) and, again consistent with our main analysis, we found that the direction of this effect was different between the south-eastern and north/western European reed warbler populations (LM; $F = 26.9$, $p < 0.0001$; see figure S8B). Finally, we found that birds were significantly closer to the expected recovery location under an inclination/bearing model of navigation than would be expected by chance (randomisation; $p < 0.001$; see methods; see figure S8C). This was not the case for any other hypothesis examined (see figure S8C).

Based on our supplementary multiple-subset analyses, we find that our result presented in the main text is repeatable using subsets of ringing data selected using multiple different (and more stringent) criteria. In turn, this implies that the results presented in the main text are robust to the precise mechanism by which breeding ringing records are parsed from migrating conspecifics and,

therefore, we suggest that it is unlikely that the inclusion of migrant birds in the analysis dataset is driving the results presented.

*Correlation matrix analysis*

Whilst our findings are consistent with the use of magnetic information during philopatry, it is possible, given the correlative nature of the study, that they, instead, result from a confound between magnetic inclination and another variable. One possible confound is environmental phenology, something to which migratory birds have been shown to be receptive (*26*). Changes in environmental phenology could cause birds to extend/truncate return migration, causing them to alter their breeding sites non-randomly and, in turn, possibly confounding changes in the Earth's magnetic field.

To test whether this was the case, we examined whether our focal magnetic parameters correlated with 4 variables related to environmental phenology: spring air temperature, the proportion of frosty days in spring, the Fraction of Absorbed Photosynthetically Active Radiation (fAPAR; a proxy for live vegetative cover) and the normalized difference vegetation index (NDVI; a second proxy for vegetative cover). All variables were derived at each position where reed warblers had been encountered, with each variable derived for as many years as possible (see methods). Each variable was normalised, and then all variables were regressed against each other in a series of uni-variate linear regressions. Because variables were normalised the gradients of different models could be directly compared, with a gradient of +1 indicating a strong positive correlation, a gradient of -1 representing a strong negative correlation and a gradient of 0 representing no correlation.

Whilst we found a substantial positive correlation between magnetic inclination and magnetic intensity (gradient = 0.31 ± 0.007), and a substantial negative correlation between magnetic inclination and magnetic declination (gradient = -0.57 ± 0.0061), we found that none of the examined environmental variables correlated notably with any magnetic variables (see table S1). The largest gradient between inclination and an environmental variable was between inclination and air temperature, with a gradient of only 0.021 (±0.0073). Based on this, we suggest that, whilst it is possible our results are the product of a confound between magnetic secular variation and another process, this seems unlikely.

| | Inclination | Declination | Intensity | fAPAR | Average air temperature | Proportion days frost | NDVI |
|---|---|---|---|---|---|---|---|
| **Inclination** | - | | | | | | |
| **Declination** | -0.57 ( ±0.0061) | - | | | | | |
| **Intensity** | 0.31 (± 0.0070) | 0.031 (± 0.0073) | - | | | | |
| **fAPAR** | -0.047 (± 0.010) | 0.019 (± 0.0099) | -0.021 (± 0.0075) | - | | | |
| **Average air temperature** | 0.021 (± 0.0073) | 0.009 (± 0.0077) | 0.048 (± 0.0073) | 0.021 (± 0.014) | - | | |
| **Proportion days frost** | -0.017 (± 0.0081) | 0.024 (± 0.0080) | -0.0057 (± 0.0087) | -0.13 (± 0.017) | 0.19 (± 0.0084) | - | |
| **NDVI** | -0.0029 (± 0.0084) | 0.008 (± 0.00709) | -0.02 (± 0.0068) | -0.053 (± 0.012) | 0.018 (± 0.0085) | 0.11 (± 0.011) | - |

**Table S1: regression matrix detailing the strength of the correlation between different magnetic parameters and potential environmental confounds.** Correlations are presented as the gradient of the regression lines between normalised variables, meaning that +1 represents a strong positive correlation, -1 a strong negative correlation and 0 represents no correlation. Gradients are presented with the 95% confidence intervals in brackets.

## Supplementary figures

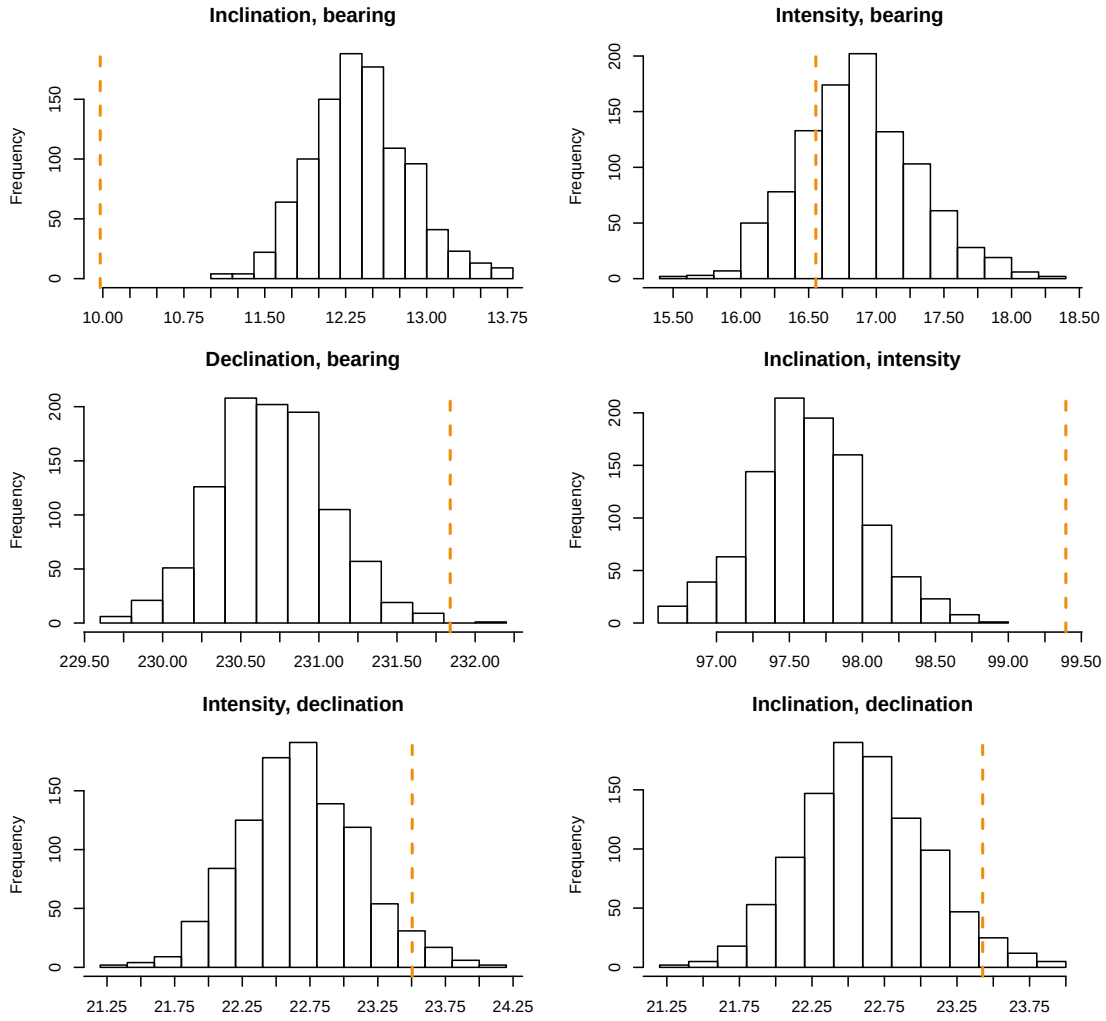



**Figure S1: observed-versus-expected analysis output for 1st year reed warblers**. The null distribution of mean 1st year observed-versus-expected distances (blue histogram) for each model of magnetic navigation compared to the observed observed-versus-expected distance (orange dotted line). The histogram is built using mean values for the observed-versus-expected distance calculated per iteration of the randomisation, and is compared to the empirically observed mean

observed-versus-expected distance. Hypotheses where the orange line is left of the histogram perform better than chance. For more information see methods and main text.

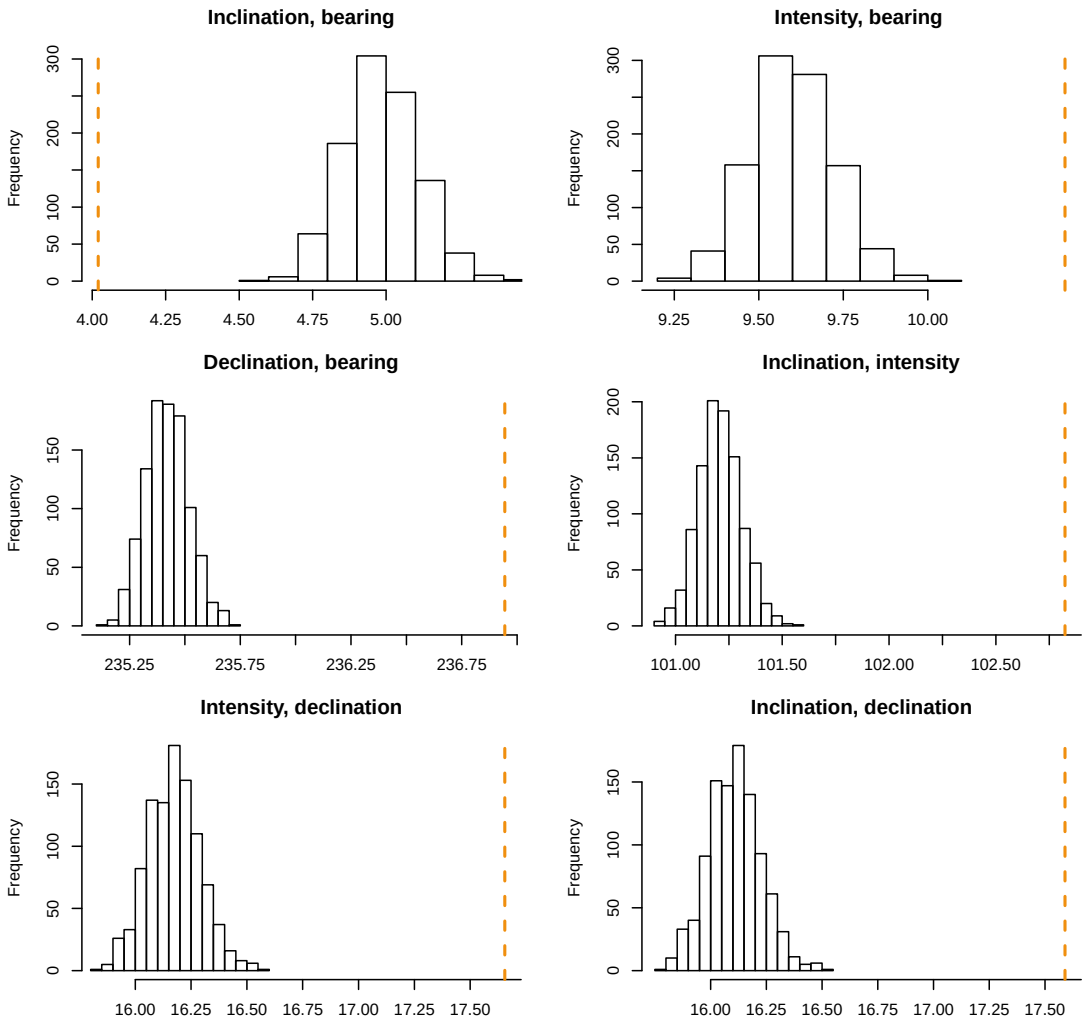
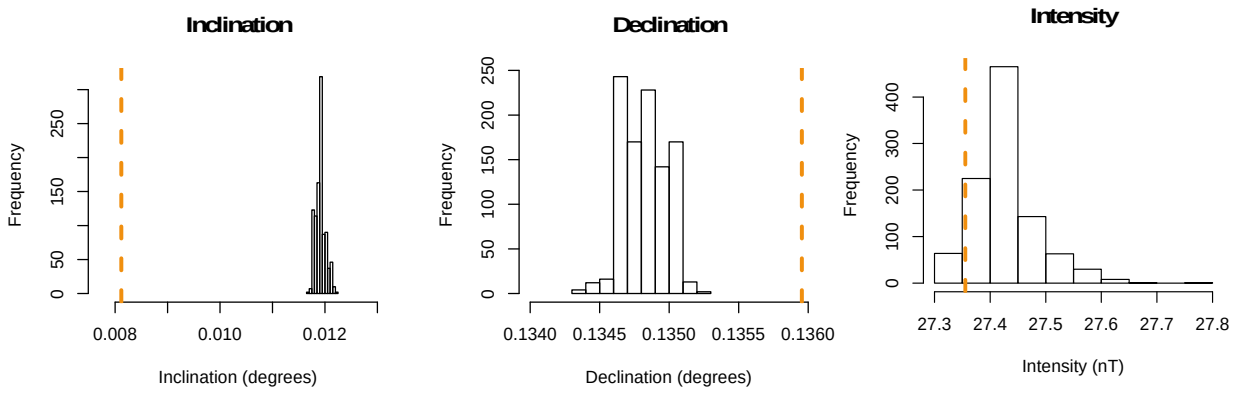
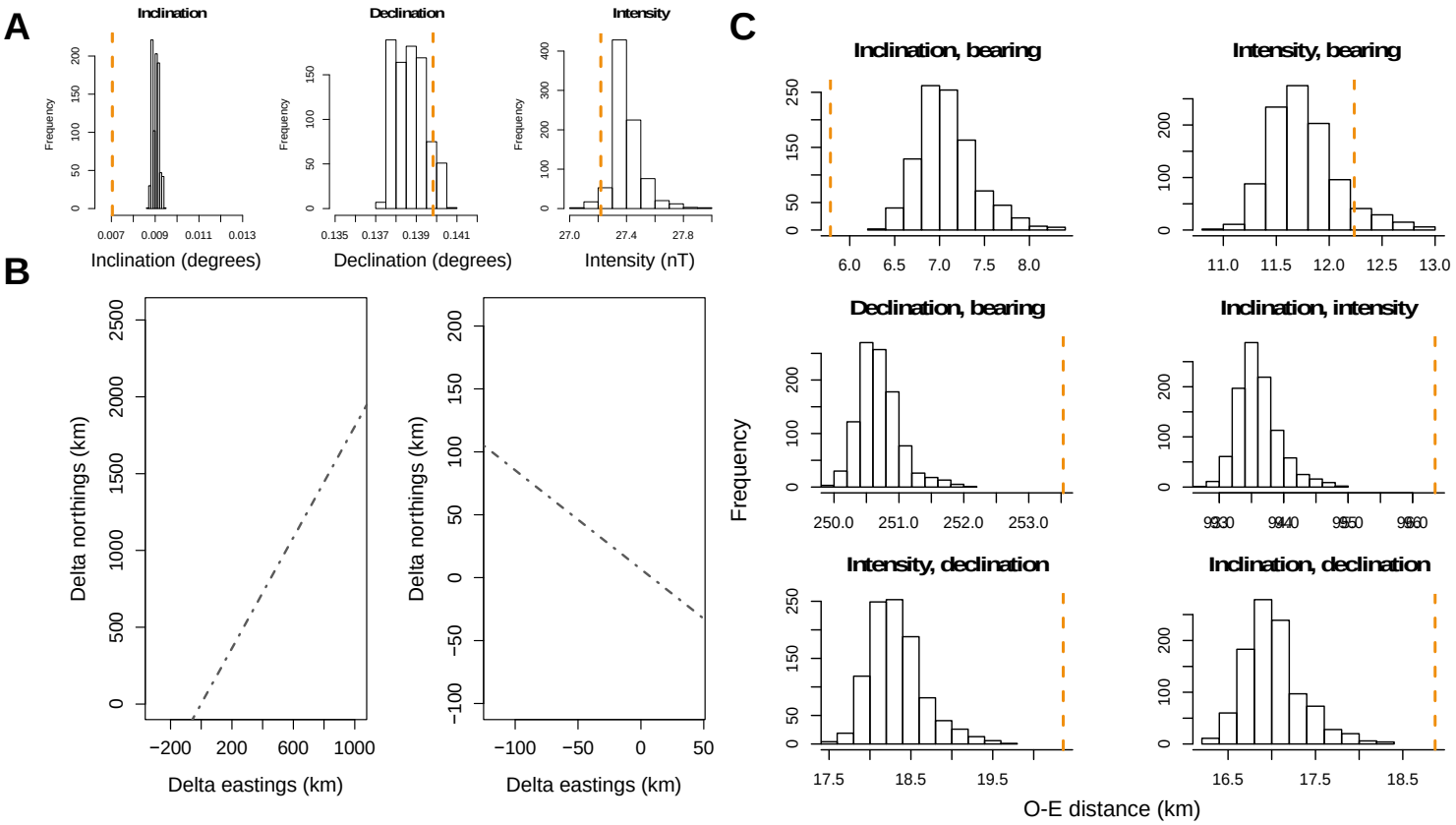



**Figure S2: observed-versus-expected analysis output for adult reed warblers**. The null distribution of mean adult observed-versus-expected distances (blue histogram) for each model of magnetic navigation compared to the observed observed-versus-expected distance (orange dotted line). The histogram is built using mean values for the observed-versus-expected distance calculated per iteration of the randomisation, and is compared to the empirically observed mean observed-versus-expected distance. Hypotheses where the orange line is left of the histogram perform better than chance. For more information see methods and main text.



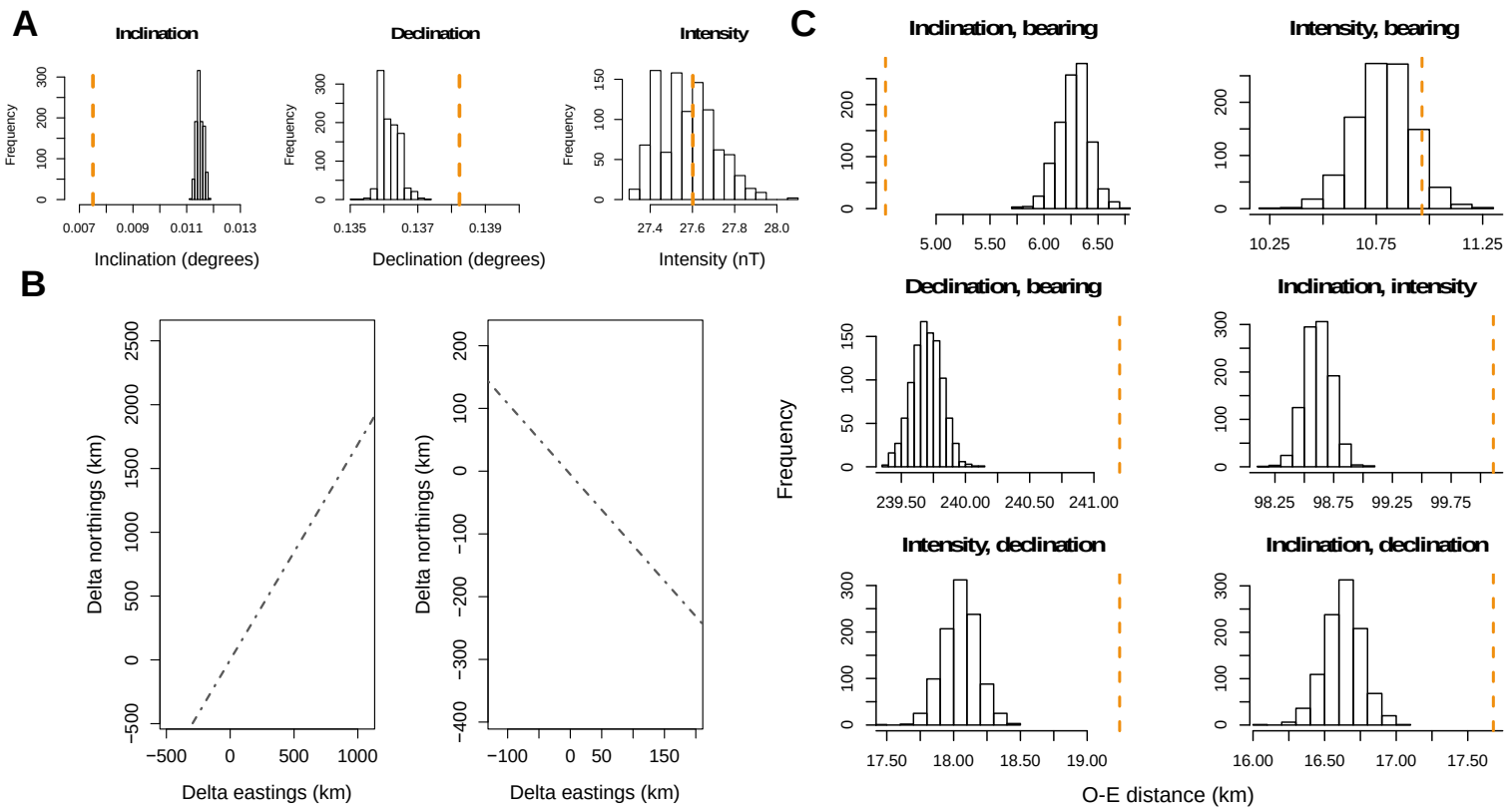

595 **Figure S3: Randomisation output comparing the observed change in magnetic parameters**
596 **to those expected if movement were random.** The median changes in each of magnetic
597 inclination (left), declination (centre) and intensity (right) are denoted by the orange dotted line.
598 The distribution of null median values for random (but equivalent) movements are denoted by the
599 blue histograms.




600

601 **Figure S4: The results of our 3 main analyses repeated for a subset of birds ringed as 1st**
602 **years at the nest and recovered as adults level with or north of the natal site.** Panel A shows
603 the results of our randomisation analysis of changes in different magnetic parameters between
604 ringing and recovery. Panel B shows the regression of changes in longitude (as 'eastings' in km)
605 against changes in latitude (as 'northings' in km) for birds breeding in north/western Europe
606 (orange) and birds breeding in south-eastern Europe (blue). In panel B point opacity represents the
607 number of overlapping datapoints. Panel C shows the outcome of our observed-versus-expected
608 randomisation for all hypotheses tested.

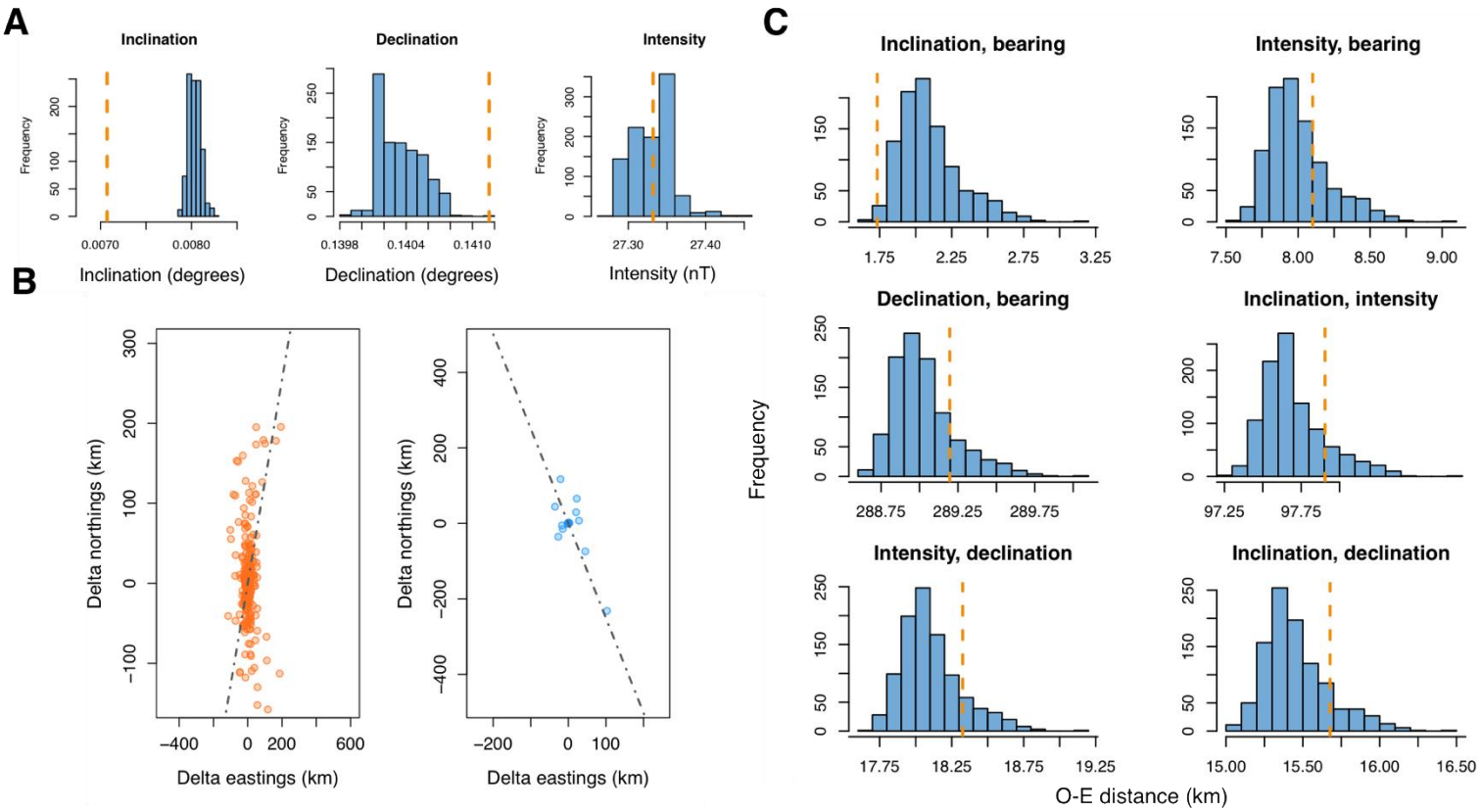
609




610

611 **Figure S5: The results of our 3 main analyses repeated for a subset of birds ringed/recovered**
612 **between 15th June and 15th July.** Panel A shows the results of our randomisation analysis of
613 changes in different magnetic parameters between ringing and recovery. Panel B shows the
614 regression of changes in longitude (as 'eastings' in km) against changes in latitude (as 'northings'
615 in km) for birds breeding in north/western Europe (orange) and birds breeding in south-eastern
616 Europe (blue). In panel B point opacity represents the number of overlapping datapoints. Panel C
617 shows the outcome of our observed-versus-expected randomisation for all hypotheses tested.

618

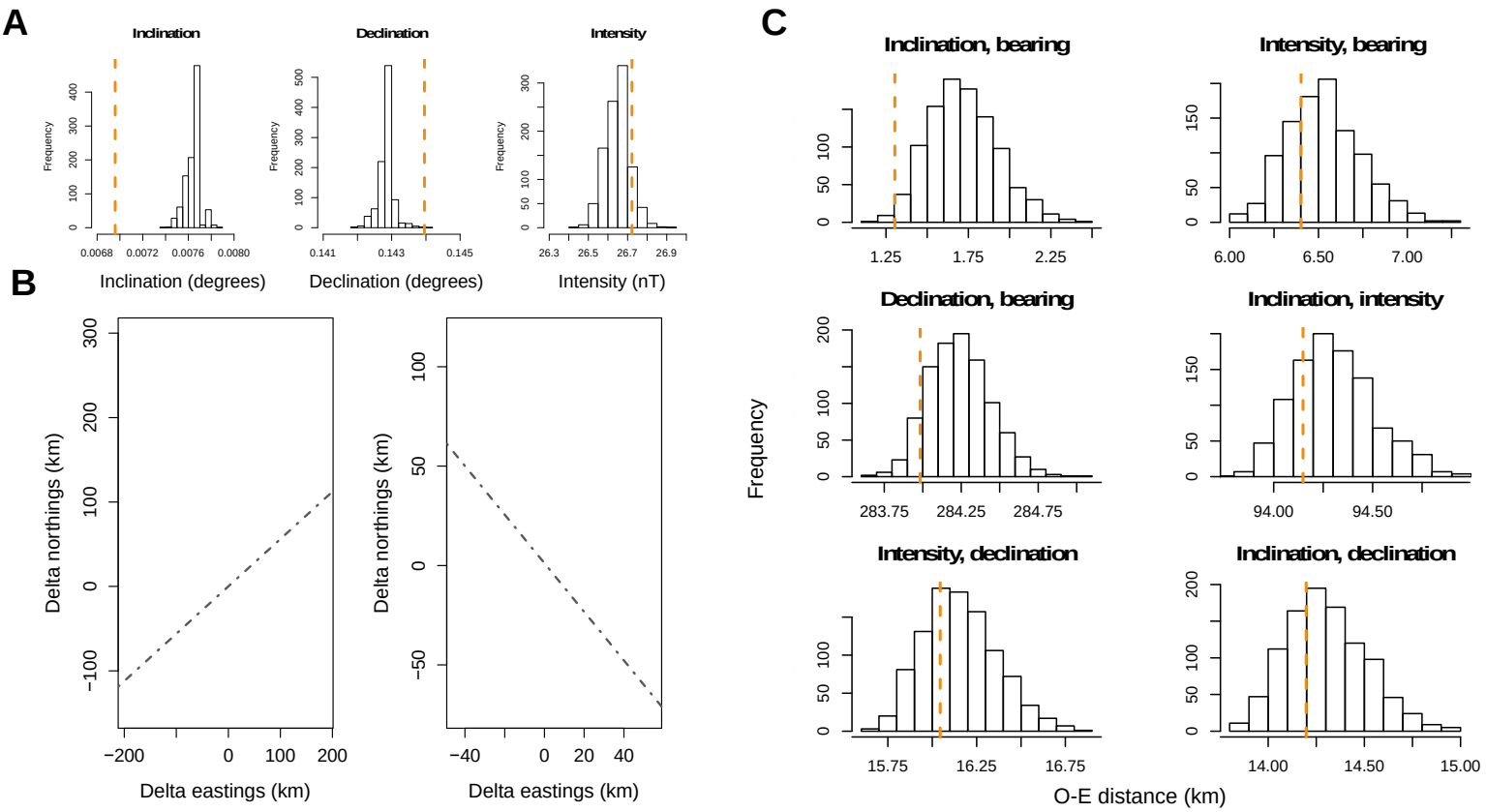



619

620 **Figure S6: The results of our 3 main analyses repeated for a subset of birds that were sexed**
621 **during both ringing and recovery.** Panel A shows the results of our randomisation analysis of
622 changes in different magnetic parameters between ringing and recovery. Panel B shows the
623 regression of changes in longitude (as 'eastings' in km) against changes in latitude (as 'northings'
624 in km) for birds breeding in north/western Europe (orange) and birds breeding in south-eastern
625 Europe (blue). In panel B point opacity represents the number of overlapping datapoints. Panel C
626 shows the outcome of our observed-versus-expected randomisation for all hypotheses tested.

627

628

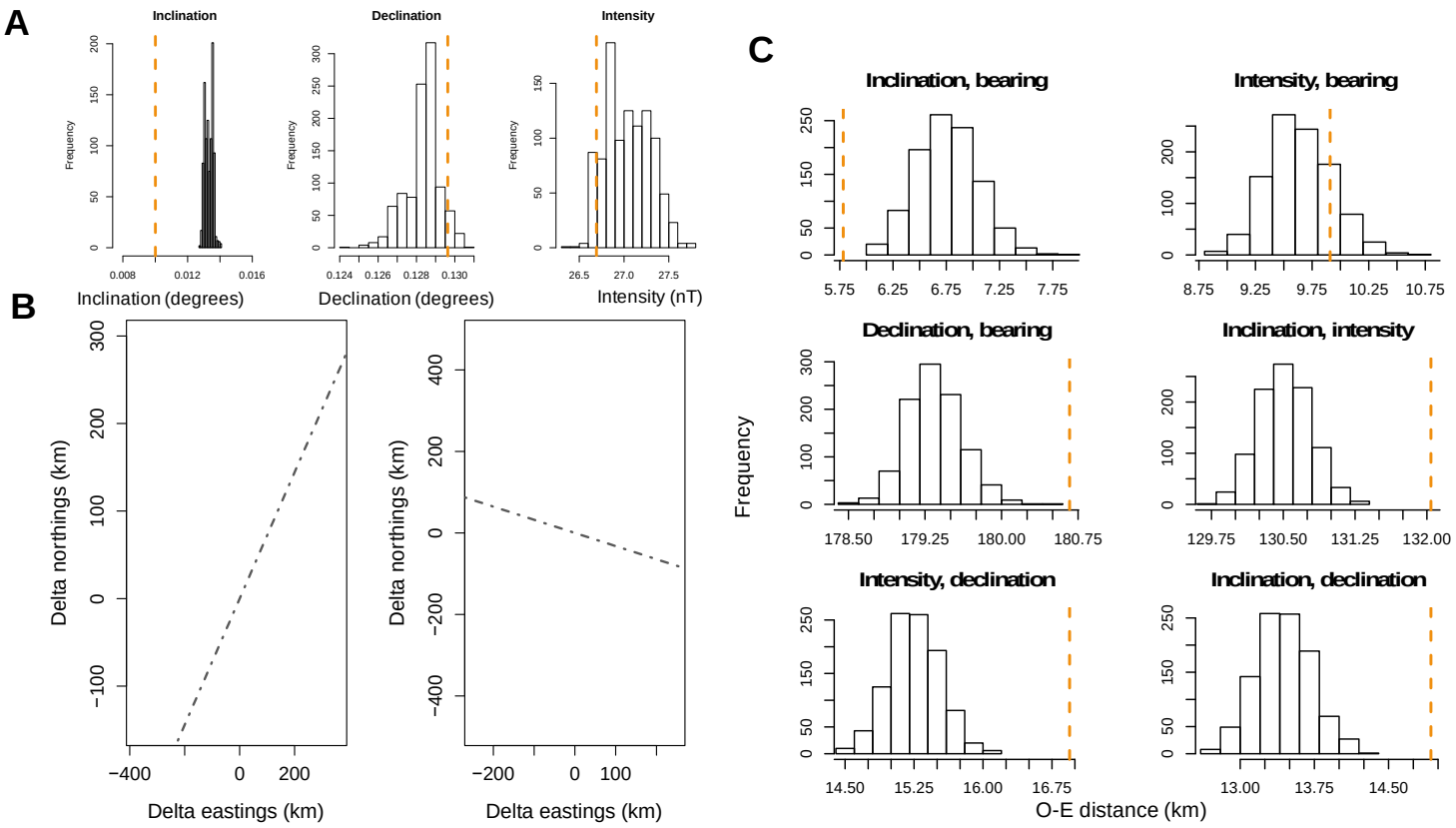



629

630 **Figure S7: The results of our 3 main analyses repeated for a subset of birds ringed after a**
631 **longitude/latitude-specified cut-off (and before the 31st July) that were sexed during both**
632 **ringing and recovery.** Panel A shows the results of our randomisation analysis of changes in
633 different magnetic parameters between ringing and recovery. Panel B shows the regression of
634 changes in longitude (as 'eastings' in km) against changes in latitude (as 'northings' in km) for
635 birds breeding in north/western Europe (orange) and birds breeding in south-eastern Europe (blue).
636 In panel B point opacity represents the number of overlapping datapoints. Panel C shows the
637 outcome of our observed-versus-expected randomisation for all hypotheses tested.

638

639




640 **Figure S8: The results of our 3 main analyses repeated for our 'constant effort' cohort (i.e.**
641 **a randomly selected group of birds where the number of recoveries per year was set at n =**
642 **50).** Panel A shows the results of our randomisation analysis of changes in different magnetic
643 parameters between ringing and recovery. Panel B shows the regression of changes in longitude
644 (as 'eastings' in km) against changes in latitude (as 'northings' in km) for birds breeding in
645 north/western Europe (orange) and birds breeding in south-eastern Europe (blue). In panel B point
646 opacity represents the number of overlapping datapoints. Panel C shows the outcome of our
647 observed-versus-expected randomisation for all hypotheses tested.
648

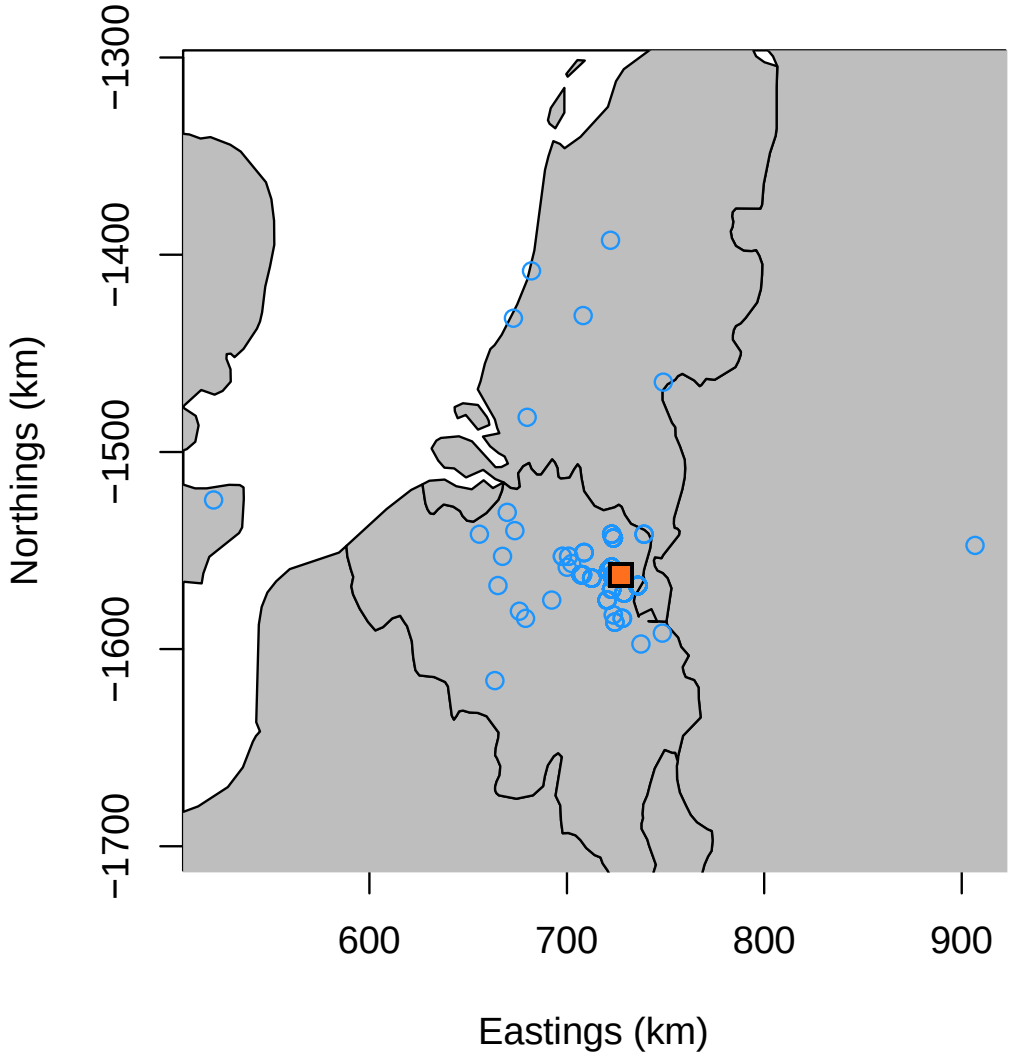



**Figure S9: The null expected recovery sites assuming that navigational error is entirely independent of secular variation.** Using our randomisation procedure (see methods), the above represents 1,000 potential recovery sites for a reed warbler breeding in northeastern Belgium. Orange/black square denotes the breeding site, translucent blue circles represent null expected recovery sites.

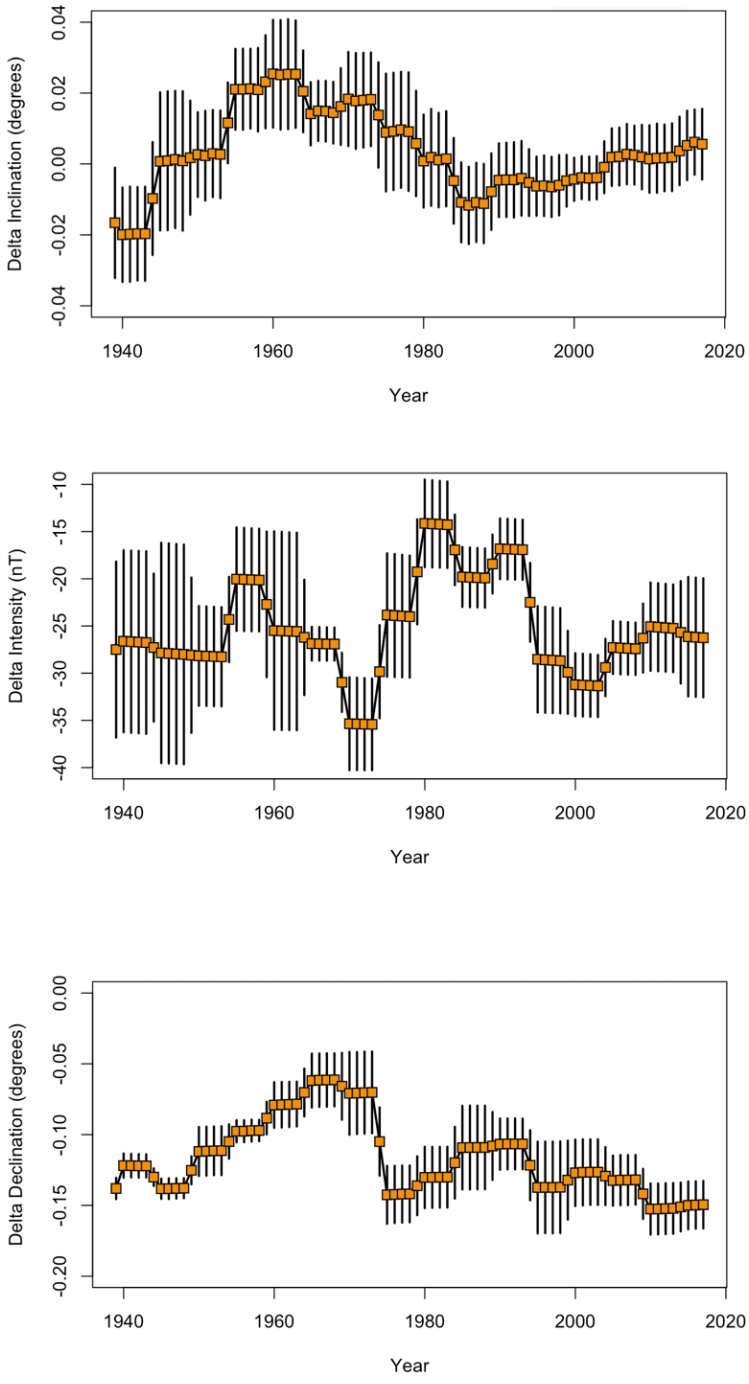



**Figure S10: Changes in magnetic parameters over the years 1940-2018.** The mean between-year shifts in magnetic inclination (in degrees), magnetic declination (in degrees) and magnetic intensity (in nanotesla) averaged from all recorded reed warbler breeding sites across Europe. Error bars show the standard deviation from across all sites.

| | Inclination | Declination | Intensity | fAPAR | Average air temperature | Proportion days frost | NDVI |
|---|---|---|---|---|---|---|---|
| **Inclination** | - | | | | | | |
| **Declination** | -0.57 ( ±0.0061) | - | | | | | |
| **Intensity** | 0.31 (± 0.0070) | 0.031 (± 0.0073) | - | | | | |
| **fAPAR** | -0.047 (± 0.010) | 0.019 (± 0.0099) | -0.021 (± 0.0075) | - | | | |
| **Average air temperature** | 0.021 (± 0.0073) | 0.009 (± 0.0077) | 0.048 (± 0.0073) | 0.021 (± 0.014) | - | | |
| **Proportion days frost** | -0.017 (± 0.0081) | 0.024 (± 0.0080) | -0.0057 (± 0.0087) | -0.13 (± 0.017) | 0.19 (± 0.0084) | - | |
| **NDVI** | -0.0029 (± 0.0084) | 0.008 (± 0.00709) | -0.02 (± 0.0068) | -0.053 (± 0.012) | 0.018 (± 0.0085) | 0.11 (± 0.011) | - |

**Table S1: regression matrix detailing the strength of the correlation between different magnetic parameters and potential environmental confounds.** Correlations are presented as the gradient of the regression lines between normalised variables, meaning that +1 represents a strong positive correlation, -1 a strong negative correlation and 0 represents no correlation. Gradients are presented with the 95% confidence intervals in brackets.

## References


1. K. Thorup *et al.*, flying on their own wings: young and adult cuckoos respond similarly to long-distance displacement during migration. *Scientific Reports* **10**, 1-8 (2020).
2. A. J. Helbig, Inheritance of migratory direction in a bird species: a cross-breeding experiment with SE-and SW-migrating blackcaps (Sylvia atricapilla). *Behavioral Ecology and Sociobiology* **28**, 9-12 (1991).
3. A. Perdeck, Two types of orientation in migrating starlings, Sturnus yulgaris L., and Chaffinches, Fringilla coelebs L., as Revealed by Displacement Experiments. *Ardea* **55**, 1-3 (1958).
4. K. Thorup *et al.*, Evidence for a navigational map stretching across the continental US in a migratory songbird. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 18115-18119 (2007).
5. R. R. Baker, *The evolutionary ecology of animal migration*. (Hodder and Stoughton, London, 1978).
6. H. Mouritsen, Long-distance navigation and magnetoreception in migratory animals. *Nature* **558**, 50-59 (2018).
7. J. R. Brothers, K. J. Lohmann, Evidence for Geomagnetic Imprinting and Magnetic Navigation in the Natal Homing of Sea Turtles. *Current Biology* **25**, 392-396 (2015).
8. N. F. Putman *et al.*, Evidence for Geomagnetic Imprinting as a Homing Mechanism in Pacific Salmon. *Current Biology* **23**, 312-316 (2013).
9. N. Chernetsov *et al.*, Migratory Eurasian Reed Warblers Can Use Magnetic Declination to Solve the Longitude Problem. *Current Biology* **27**, 2647-+ (2017).
10. N. Chernetsov, A. Pakhomov, A. Davydov, F. Cellarius, H. Mouritsen, No evidence for the use of magnetic declination for migratory navigation in two songbird species. *Plos One* **15**, (2020).
11. B. Leisler, K. Schulze-Hagen, D. Quinn, Max Planck Institute for Ornithology., *The reed warblers : diversity in a uniform bird family*. (KNNV, Zeist, 2011), pp. 327 p.
12. R. Tittler, M. A. Villard, L. Fahrig, How far do songbirds disperse? *Ecography* **32**, 1051-1061 (2009).
13. R. A. Holland, True navigation in birds: from quantum physics to global migration. *Journal of Zoology* **293**, 1-15 (2014).
14. H. Mouritsen, *Avian Migration*. (Springer Verlag, Berlin, 2003).
15. N. Chernetsov, D. Kishkinev, S. Gashkov, V. Kosarev, C. V. Bolshakov, Migratory programme of juvenile pied flycatchers, Ficedula hypoleuca, from Siberia implies a detour around Central Asia. *Animal Behaviour* **75**, 539-545 (2008).
16. E. Gwinner, W. Wiltschko, ENDOGENOUSLY CONTROLLED CHANGES IN MIGRATORY DIRECTION OF GARDEN WARBLER, SYLVIA-BORIN. *Journal of Comparative Physiology* **125**, 267-273 (1978).
17. P. Prochazka *et al.*, Across a migratory divide: divergent migration directions and non-breeding grounds of Eurasian reed warblers revealed by geolocators and stable isotopes. *Journal of Avian Biology* **49**, (2018).
18. L. Zwarts, R. G. Bijlsma, J. Van der Kamp, E. Wymenga, *Living on the edge: wetlands and birds in a changing Sahel*. (BRILL, 2012).

19. L. Svensson, *Identification guide to European passerines*. (Naturhistoriska Riksmuseet, Stockholm, 1970), pp. 152 p.
20. J. R. Brothers, K. J. Lohmann, Evidence that Magnetic Navigation and Geomagnetic Imprinting Shape Spatial Genetic Variation in Sea Turtles. *Current Biology* **28**, 1325-+ (2018).
21. J. Wynn, O. Padget, H. Mouritsen, C. Perrins, T. Guilford, Natal imprinting to the Earth's magnetic field in a pelagic seabird. *Current biology : CB*, (2020).
22. L. Halupka, A. Dyrcz, M. Borowiec, Climate change affects breeding of reed warblers Acrocephalus scirpaceus. *Journal of Avian Biology* **39**, 95-100 (2008).
23. N. Chernetsov, Timing of spring migration, body condition, and fat score in local and passage populations of the Reed Warbler Acrocephalus scirpaceus on the Courish Spit. *Avian Ecol Behav* **2**, 75-88 (1999).
24. R. C. Team. (2017).
25. I. Newton, K. Brockie, *The migration ecology of birds*. (Elsevier-Academic Press, Amsterdam ; London, 2008), pp. viii, 976 p.
26. K. Thorup *et al.*, Resource tracking within and across continents in long-distance bird migrants. *Science Advances* **3**, (2017).