1   **Title: Attributing changes in the distribution of species abundance to weather variables**

2   **using the example of British breeding birds**

3

4   **Short Running Title: Attributing abundance changes to weather variables**

5   Cornelia S. Oedekoven[1*], David A. Elston[2], Philip J. Harrison[1], Mark J. Brewer[2], Stephen T.

6   Buckland[1], Alison Johnston[3], Simon Foster[4] and James W. Pearce-Higgins[3]

7

8   [1]*Centre for Research into Ecological and Environmental Modelling, University of St.*

9   *Andrews, The Observatory, St Andrews, Fife, KY16 9LZ, UK;*

10   [2]*Biomathematics & Statistics Scotland, Craigiebuckler, Aberdeen, AB15 8QH, UK;*

11   [3]*British Trust for Ornithology, The Nunnery, Thetford, Norfolk, IP24 2PU, UK*

12   [4]*Scottish Natural Heritage, Great Glen House, Leachkin Road, Inverness IV3 5NW, UK*

13   ** Corresponding author. Email: cso2@st-and.ac.uk*

14

15   **Word count:** 6999

16   Summary: 337

17   Main body (Summary to Discussion): 4990

18   References: 1041

19   Acknowledgments and Data Accessibility: 96

20   Tables & Figures: 872

21

Summary

22

23 **1.**   Modelling spatio-temporal changes in species abundance and attributing those

24 changes to potential drivers such as climate, is an important but difficult problem. The

25 standard approach for incorporating climatic variables into such models is to include each

26 weather variable as a single covariate whose effect is expressed through a low-order

27 polynomial or smoother in an additive model. This, however, confounds the spatial and

28 temporal effects of the covariates.

29 **2.**   We developed a novel approach to distinguish between three types of change in any

30 particular weather covariate. We decomposed the weather covariate into three new covariates

31 by separating out temporal variation in weather (averaging over space), spatial variation in

32 weather (averaging over years) and a space-time anomaly term (residual variation). These

33 three covariates were each fitted separately in the models. We illustrate the approach using

34 generalized additive models applied to count data for a selection of species from the UK's

35 Breeding Bird Survey, 1994-2013. The weather covariates considered were the mean

36 temperatures during the preceding winter and temperatures and rainfall during the preceding

37 breeding season. We compare models that include these covariates directly with models

38 including decomposed components of the same covariates, considering both linear and

39 smooth relationships.

40 **3.**   The lowest QAIC values were always associated with a decomposed weather

41 covariate model. Different relationships between counts and the three new covariates

42 provided strong evidence that the effects of changes in covariate values depended on whether

43 changes took place in space, in time, or in the space-time anomaly. These results promote

44 caution in predicting species distribution and abundance in future climate, based on

45 relationships that are largely determined by environmental variation over space.

46   **4.**   Our methods estimate the effect of temporal changes in weather, whilst accounting for

47   spatial effects of long-term climate, improving inference on overall and/or localised effects of

48   climate change. With increasing availability of large-scale data sets, need is growing for

49   appropriate analytical tools. The proposed decomposition of the weather variables represents

50   an important advance by eliminating the confounding issue often inherent in large-scale data

51   sets.

52

53   **Key words:** *climate change; decomposition of spatial, temporal and anomaly effects;*

54   *generalized additive models; generalized linear models; spatio-temporal modelling; species*

55   *abundance; UKCP09 climate projections.*

56

57   **1. Introduction**

58   It has been demonstrated for a wide range of taxa that climate change has an effect on

59   abundance and distribution of individual species as well as community measures of

60   biodiversity (e.g. Talley, Coley & Kursar 2002; Schummer et al. 2010; WallisDeVries,

61   Baxter & Van Vliet 2011). In order to detect or predict the impacts of climate change on

62   biodiversity, many studies either analyse spatial variation in species' abundance or

63   occurrence as a function of spatial variation in climate (e.g. Beale et al. 2013, Renwick et al.

64   2012, Johnston et al. 2013), or temporal variation in abundance in relation to temporal

65   changes in weather variables (e.g. Davey et al. 2012, Devictor et al. 2012, Pearce-Higgins et

66   al. 2015). However, the potential to improve inference from combined spatio-temporal

67   variation in both biological and climate/weather variables has rarely been considered. The

68   standard approach currently available for this is to include either a single covariate in a

69   generalized linear model, or a single smoothing term in a generalized additive model (GAM)

70   (e.g. Araújo et al. 2005). However, this approach confounds the spatial and temporal effects

3

71   of the covariates. This confounding means, e.g., that a change in temperature in one location

72   from 12°C to 13°C would have the same effect as the difference between two locations at one

73   point in time: one location at 12°C and one at 13°C.

74   We developed a method with more flexibility in attributing patterns in abundance

75   which disentangles the temporal and spatial effects of the weather covariates via

76   decomposition into three components. The temporal term is the average of observed covariate

77   values over space for a given time; the spatial term is the average over time for a given

78   location; the space-time anomaly term is the residual variation for a given location and time.

79   When using these decomposed covariates, a difference in the original covariate can be

80   associated with the temporal, spatial or residual variation component or a combination of the

81   components. This method allows the fitted models to differentiate between spatial, temporal

82   and spatio-temporal variation in the original weather covariate on abundance, thus alleviating

83   the confounding issue.

84   To investigate changes in species abundances over space and time, data are required

85   from well-designed long-term surveys such as the UK Breeding Bird Survey (BBS). The UK

86   BBS is a large-scale biodiversity monitoring programme with a protocol that allows spatio-

87   temporal correlation in animal abundances *and* differences in detectability to be accounted

88   for (Yoccoz, Nichols & Boulinier 2001; Risely et al. 2013), both of which can cause serious

89   biases in analyses of biodiversity or abundance trends (Buckland et al. 2012). To

90   accommodate imperfect detectability in observed counts, sites are surveyed using line-

91   transect methods, an example of distance sampling (Buckland et al. 2001). We accounted for

92   spatial and temporal autocorrelation in the observed counts using modelling techniques for

93   abundance estimation that smooth across space and time (Harrison et al. 2014). Besides

94   climate, the spatial distribution and abundance of bird species are affected by other factors

95   defining habitat suitability (Gregory & Baillie 1998); hence, we also included land class and

96 elevation as covariates. To illustrate our methods, we selected three weather covariates for

97 our models based on Eglington & Pearce-Higgins (2012): the centred means of temperatures

98 during the preceding winter and temperatures and rainfall in the preceding breeding season.

99   We describe the decomposition of the weather covariates as well as the modelling

100 approach for the BBS data (Section 2) and compare models that include these covariates

101 directly with models including decomposed components of the same covariates. We consider

102 linear (Section 3.1) and smooth terms (flexible, nonlinear model components, Sections 3.2,

103 3.4) for these covariates to assess the need for fitting smooth terms for weather covariates

104 rather than linear terms. We further investigate the necessity of a space-time smooth in the

105 model after fitting the weather covariates as smooth effects (Section 3.3). The best model is

106 used to predict abundances for each species (Section 3.5). Lastly, we discuss the benefits of

107 our approach for biological interpretations of the models (Section 4).

108

109 **2. Methods**

110 2.1 DATA

111 We begin by describing the data for the response variable of the models, the BBS bird counts,

112 which determines the spatial and temporal resolution required for the weather data. We then

113 describe our newly proposed method of decomposing the weather variables in Section 2.1.2.

114 *2.1.1 BBS data*

115 Sampling sites for the BBS are randomly selected 1km squares; the number of sites surveyed

116 each year has increased since the survey began, from ~1500 in 1994 to 3350 in 2013. Sites

117 were visited twice per year (April to early-May and late-May to June), during which

118 volunteers walked two parallel 1km transect lines and assigned each detected bird to one of

119 four categories (0-25m from the line, 25-100m, >100m, flying). We only considered data

120 from the first two distance intervals for which detection probabilities were adequate (>0.1)

121   for model fitting and estimating average detection probabilities (the proportion of birds

122   detected within the surveyed area, Buckland et al. 2001). We minimised the possibility of

123   including juvenile birds by using data only from the early visit for most species, and the later

124   visit for late breeding birds such as summer migrants (Table 1).

125   We analysed the BBS data collected in 1994-2013, excluding 2001 when an outbreak

126   of foot-and-mouth disease restricted access to many areas (Risely et al. 2013), and present

127   results for five species of birds that are likely to show a range of sensitivities to changes in

128   climate, namely goldcrest (*Regulus regulus*), song thrush (*Turdus philomelos*), linnet

129   (*Carduelis cannabina*), cuckoo (*Cuculus canorus*) and willow warbler (*Phylloscopus*

130   *trochilus*). This choice of species included residents whose populations may have increased

131   in response to recent warming (goldcrest, song thrush and linnet, Pearce-Higgins et al. 2015),

132   and declining long-distance migrants whose populations may be affected by conditions in the

133   UK, during migration or in Africa (cuckoo and willow warbler, Ockendon, Johnston &

134   Baillie 2014).

135   A complication for spatial modelling is that some species are rare or absent in parts of

136   Britain; if we extrapolated our predictions for such species to those areas, the predictions

137   might be unstable and not meaningful. Therefore, we created a grid of 100km squares

138   throughout Britain and restricted our modelling for each species to those 100km squares for

139   which there were a total of over three positive observations of the species during the years of

140   the study (Harrison et al. 2014; "propM" from Table 1 gives the proportion of Britain

141   included in the models by species).

142

143   *2.1.2 Weather data and decomposition method*

144   For the weather data 1993-2013 we used UKCP09 5x5km gridded observation data compiled

145   by the Met Office (www.metoffice.gov.uk). To model abundance in a given breeding season,

146 we included three weather variables for the preceding winter and preceding breeding season.

147 To reduce correlations between parameter estimates we used their centred values which for a

148 given location and time point is equal to the original value minus the mean over all locations

149 across Britain and time points for the respective covariate. We indexed the values according

150 to the year of the corresponding response variable:

151 $W_{i,t}$ = centred monthly means of the daily minimum temperature during the preceding winter

152 (Dec-Feb) in year $t - 1$ (Dec) and t (Jan-Feb) at location i;

153 $B_{i,t}$ = centred monthly means of the daily mean temperature during the preceding breeding

154 season (April-July in year $t - 1$) at location i;

155 $P_{i,t}$ = centred square root of the monthly means of the total rainfall during the preceding

156 breeding season (April-July in year $t - 1$) at location i. We use the square root to reduce the

157 effect of very large rainfall values.

158 We note that the 5km resolution weather variable values get applied to each 1km

159 square within them. We refer to the centred weather covariates as W, B and P (short for

160 winter, breeding and precipitation) for brevity and decompose each of them into three

161 components: the average over space for a given time (year); the average over time for a given

162 spatial location and; the residual variation term. For example, for winter minimum

163 temperature we define the first component, the *average over space for a given time* (temporal

164 component), as:

165 $W_{time}: W_{\cdot,t} = \dfrac{\sum_i W_{i,t}}{N}$

166 where N is the number of 5km grids across Britain. The second component, the *average over*

167 *time for a given spatial location* (spatial component) is given by:

168 $W_{space}: W_{i,\cdot} = \dfrac{\sum_t W_{i,t}}{T}$

169 where T gives the number of time points. Finally the third component, the *residual variation*

170 *term* (or *space-time anomaly*), which can be thought of as the difference between the original

171 value and what is expected given the overall yearly and overall spatial effect, is defined as:

172 W_{resid}: $W'_{i,t} = W_{i,t} - W_{\cdot,t} - W_{i,\cdot}$

173   For any of the three centred covariates W, B and P, we distinguish between the three

174 corresponding decomposed variables using the subscripts *time*, *space* or *resid*, e.g., where

175 W_{time} denotes the average over space for a given time for the preceding winter temperatures

176 (W).

177   Figures 1-3 graphically depict the decomposed variables for W, B and P, showing the

178 patterns of variation between years averaging over space, between grid cells averaging over

179 years, and the residual variation.

180

181 *2.1.3 Land class and elevation data*

182 In addition to terms describing the weather variables of interest, it is beneficial to include

183 other terms in the model to remove some aspects of broad-scale variation related to other

184 factors. Following Harrison et al. (2016), we include elevation and habitat information.

185 Habitat information was obtained from the 2007 land cover map compiled by the Centre for

186 Ecology and Hydrology (CEH). This gives the percentages of 10 aggregate land classes

187 (broadleaf woodland, coniferous woodland, arable, improved grassland, semi-natural

188 grassland, mountain/heath/bog, saltwater, freshwater, coastal and built-up areas/gardens) for

189 all 1km squares across Britain(https://catalogue.ceh.ac.uk/documents/c3723adb-1a8c-4b57-

190 958b-1d610d2c37fe, Morton et al. 2011).

191

192 2.2 DATA ANALYSES

193 *2.2.1 Models for the relationships between weather and species abundance*

194   For each of the five species, we fitted GAMs assuming a quasi-Poisson distribution (Section

195   2.2.2) with a log-link function to the count data from each recorded site and year. All models

196   contained elevation and land class covariates, plus a selection amongst a space-time smoother

197   and the weather covariates which were of particular interest. See below and Table 2 for

198   model summaries and explanation of model terms.

199   Overall, these models can be expressed as:

200    $E[Y_{s,i,t}] = \exp\left(\alpha_s + \beta_{e,s}elev_i + \sum_{l=1}^{L}\beta_{l,s}pac_{l,i} + f_s(east_i, north_i, t) + \mathcal{M}_{s,i,t}\right)$ (1),

201   where $E[Y_{s,i,t}]$ gives the expected count for species s on 1km square i in year t; $elev$ the

202   mean elevation; pac_l the percentage of each of the $L (= 10)$ land classes; f_s is a space-time

203   smooth and $\mathcal{M}_{s,i,t}$ pertains to the weather covariates and varied between models. The location

204   of squares was given by eastings ($east$) and northings ($north$). GAMs were fitted using

205   package mgcv (version 1.8-4, Wood 2006) in the statistical software R (R core development

206   team 2011).

207   For the space-time smooth, f_s, we used a tensor product of a thin plate regression

208   spline (TPRS) of $east$ and $north$ and a TPRS of t (Wood 2006, p. 225). Models thus

209   incorporate interactions between space and time and the spatial component captures, along

210   with any spatial autocorrelation, differences in the character of the land cover across Britain

211   (Renwick et al. 2012). The amount of smoothing is described by the effective degrees of

212   freedom (EDF). In accordance with our previous analysis of biodiversity trends in breeding

213   birds (Harrison et al. 2014) we set an upper limit of five for the EDF of the temporal TPRS,

214   with the actual value being determined by in-built cross-validation. By trial and error we

215   found 25 to be a suitable upper limit for the EDF of the spatial TPRS. If, however, some of

216   the 100km grids had been removed for the species (Section 2.1.1), the maximum EDF for the

217   spatial TPRS was scaled according to the proportion of Britain remaining ("propM", Table

218   1). To minimise unwanted edge effects (unrealistic relationships for smooth terms in extreme

219   regions of the covariate) when modelling the cuckoo, for which 96% of counts were zero, we

220   divided the upper limits of both EDFs by two.

221   We used the model without any weather covariates, i.e. the term from eqn (1)

222   pertaining to the weather covariates, $\mathcal{M}_{s,i,t}$ equals zero, as a baseline model *mSTS*. We

223   investigated the effects of the weather variables firstly by considering them as linear

224   covariates, with the further aim of disentangling their spatial, temporal and residual

225   components. Building on *mSTS*, we fitted seven models with at least one linear effect of the

226   weather covariates (Table 2A). In these models, $\mathcal{M}_{s,i,t}$ contained linear functions of either

227   only a single centred weather covariate (i.e. not decomposed into three components), all three

228   centred weather covariates or two centred and one decomposed weather covariate. We

229   refrained from decomposing more than one weather covariate in a single model to avoid

230   overfitting and complexity of model interpretation.

231   As linear effects may not capture the relationship sufficiently, we replaced these with

232   smooth functions of the centred and decomposed weather covariates to create seven further

233   models, each corresponding to a smooth equivalent of one of the models with linear weather

234   terms (Table 2B). Smooth functions allow capturing nonlinear relationships between counts

235   and covariates (on the linear predictor scale) using the flexibility of splines.

236   We investigated the consequences of removing the space-time smooth from the model

237   because the need for this term may be reduced by including covariates that are varying in

238   space and time and because the presence of this term may be influencing the estimated form

239   of the weather variable effects. Here we used the seven models with the smooth functions for

240   the weather covariates and fitted the corresponding models without the space-time smooth for

241   comparison, i.e. omitting f_s from the model (Table 2C). Model *m_* corresponds to *mSTS*

242   without the space-time smooth.

243   For each species, the overall best model was used to predict abundances throughout

244   Britain (excluding areas not part of the analysis for that species) using the 1km squares from

245   Section 2.1.3. Here, we used the estimates of average detection probability for each species

246   (Section 2.1.1) to scale up from predicted counts to predicted abundance within the 1km

247   squares.

248

249   *2.2.2 Model comparisons*

250   In all the models we assumed a quasi-Poisson error structure for the $Y_{s,i,t}$ (observed counts) to

251   allow for over- or underdispersion. Overdispersion is a common feature when simple

252   statistical models are fitted to count data due to heterogeneity that the models have failed to

253   account for. Underdispersion can occur if birds are territorial and hence self-organising in a

254   manner that can lead to less-than-random variation. In the standard Poisson model the

255   residual variance equals the mean. For the quasi-Poisson model, we relax this constraint to

256   proportionality, thus moving from "*variance=mean*" to "*variance=θ×mean*" in which the

257   scale parameter $θ$ determining the variance-mean relationship is estimated during the model

258   fitting ($\hat{θ} > 1$ corresponding to overdispersion, $\hat{θ} < 1$ to underdispersion) We used deviance

259   residuals for estimating $θ$ for more reliable variance estimation compared to the default

260   Pearson residuals (Harrison et al. 2016). Standard errors for parameters of a quasi-Poisson

261   model are adjusted by multiplying the standard errors of the equivalent Poisson model with

262    $\sqrt{\hat{θ}}$.

263   We used QAIC values (Richards 2008) for model comparison. QAIC differs from

264   AIC in that the log-likelihood component is divided by $\hat{θ}_c$ which is the $\hat{θ}$ from the most

265   complex of the candidate models. Three models were equally complex – those with the

266   space-time smoother and three weather covariates of which one was decomposed. For a given

267   species, we used the smallest $\hat{θ}$ from these three models as $\hat{θ}_c$ for all models (which also

11

268   happened to be the smallest $\hat{\theta}$ between all candidate models). For cuckoo where $\hat{\theta} < 1$ for all

269   models, we set $\hat{\theta}_c = 1$. We describe differences between QAIC values as δ-QAIC, which

270   represent differences from the model with the smallest QAIC value for each species.

271

272   **3. Results**

273   In the following, we compare the linear and smooth functions of the weather variables from

274   the different models by illustrating the coefficients and smooth functions for the respective

275   covariate. Model definitions and δ-QAIC values are given in Table 2.

276

277   *3.1 Including weather covariates as linear effects in the models*

278   For all species except cuckoo, the best of the models with linear covariates used a

279   decomposed covariate, *W* for goldcrest, song thrush, linnet; *P* for willow warbler.

280   Inspection of the regression coefficients (Figure 4) showed that those for the three

281   decomposed covariates corresponding to a single weather variable can be of very different

282   magnitude and, most notably for goldcrest *mSTS_lWPBd*, of different signs. Where one of the

283   decomposed covariate values differed substantially from the others, it was always the spatial

284   decomposition that stood apart from the temporal and anomaly coefficients.

285

286   *3.2 Replacing linear terms with smooth effects*

287   Using smooth instead of linear effects generally led to moderate or substantial improvement

288   of model fit (compare e.g. Table 2B row *mSTS_sW* vs the corresponding entry in Table 2A

289   rows *mSTS_lW* or 2B row *mSTS_sWBPd* vs 2A row *mSTS_lWBPd*).

290   The smoothers associated with each model for each species are shown in Appendix 1; we

291   show a selection of smoothers in Figure 5. For these models with the space-time smoother,

292   the influence of any weather variable on any given bird species depended very little on

293   whether the other weather covariates were included in the model or not (e.g. for goldcrest,

294   compare the W smoothers for models $mSTS\_sW$ and $mSTS\_sWBP$. Furthermore, when any

295   one weather variable was decomposed, the estimated effects of the other weather variables

296   changed little (e.g. for goldcrest, compare the smoothers for W between the models

297    $mSTS\_sWBP$, $mSTS\_sWBPd$ and $mSTS\_sWPBd$).

298

299   *3.3 Dropping the space-time smoother*

300   Dropping the space-time smooth, either from the models without any weather covariates, or

301   the models with smooth weather covariates, led to enormous increases in QAIC (Table 2, $m\_$

302   vs $mSTS$ and Table 2C vs the corresponding elements of Table 2B). However, whether the

303   space-time smoother was included in the model had little effect on conclusions regarding

304   weather covariates apart from a few instances (e.g. willow warbler, B).

305

306   *3.4 Descriptions of best models, all containing decomposed smoothed terms*

307   For goldcrest, the best model $mSTS\_sBPWd$ suggests abundance was higher in colder

308   locations (W_{space}) and in warmer winters (W_{time}) and when and where there were positive

309   anomalies (W_{resid}) (Figure 5). All three of these smoothers differed markedly in shape from

310   the single smoother for W prior to decomposition in model $mSTS\_sWBP$, demonstrating the

311   decomposition is able to describe different functional responses when separated.

312   For song thrush, the best model $mSTS\_sWPBd$ suggests little effect on abundances of

313   temperature in the preceding breeding season (B_{time}) or the anomalies (B_{resid}) (Figure 5). The

314   smoother for B_{space} (the spatial covariate) is similar in shape to the smoother for B prior to

315   decomposition, but with a more marked effect.

316   For linnet, the best model $mSTS\_sBPWd$ indicates little effects of W_{time} or W_{resid} but

317   lower abundances in locations with cold winters (W_{space}) (Figure 5). Thus the mostly positive

13

318   relationship of the single covariate W is almost entirely due to spatial variation in mean

319   winter temperatures (as opposed to variation over time or the anomaly).

320   For cuckoo the best model $mSTS\_sWBPd$ indicates higher abundance in locations

321   which generally have drier breeding seasons (P_{space}) and less residual variation in

322   precipitation (P_{resid}) while the temporal covariate (P_{time}) showed little effect (Figure 5). This

323   pattern was only evident when decomposing covariate P.

324   For willow warbler, the best model $mSTS\_sWPBd$ revealed higher abundances for

325   locations with moderate breeding season temperatures (B_{space}) and negligible effects of the

326   temporal and anomaly components (B_{time} and B_{resid}) (Figure 5). The pattern from the

327   undecomposed B was similar to the spatial effect and therefore largely represents the spatial

328   component of breeding season temperature.

329

330   *3.4 Abundance predictions*

331   We used the best model for each species to construct maps of predicted abundances

332   throughout Britain for each year 1994-2013 (Appendix 2 shows all years); we present four

333   selected years, roughly equally separated: 1994, 2000, 2006 and 2013 (Figure 6). For each

334   species, the best model contained a decomposed weather covariate for which the spatial

335   component showed the strongest pattern (Figure 5). The relationship between abundance and

336   elevation was negative for all species, whereas the relationships between numbers and the ten

337   land classes varied between species.

338

339   **4. Discussion**

340   There is an increasing recognition of the potential to analyse long-term biodiversity

341   monitoring data sets to document the impacts of long-term environmental changes upon

342   species' distributions, abundances and communities. Interest in the consequences of climate

343 change has resulted in a significant increase in the number and types of these studies. These

344 often either collate data from multiple locations to identify temporal trends and analyse those

345 in relation to climate change (e.g. Davey et al. 2012, Devictor et al. 2012, Pearce-Higgins et

346 al. 2015), or model spatial variation in abundance or occurrence as a function of climate to

347 describe changes in distribution (e.g. Beale et al. 2013, Renwick et al. 2012, Johnston et al.

348 2013). Often sufficient resolution in time and space is not available in a dataset, so spatial

349 variation is used implicitly or explicitly as a proxy for temporal variation (e.g. Chen et al.

350 2011, Bellard et al. 2012 and Warren et al. 2013), where, e.g. the estimated species response

351 between cold and warm locations is used to estimate the species response in cold locations

352 that warm under climatic change. We proposed a novel way of including weather variables in

353 spatio-temporal models of abundance that involves decomposing each weather variable into a

354 spatial component, a temporal component and anomalies. Even after fitting a space-time

355 smoother alongside land class and elevation covariates in our models, we found that for all

356 five species the best models involved the decomposed covariates. This may have been due to

357 the divergence of spatial and temporal responses to the covariates.

358   Our methods enabled us to obtain improved understanding of how these effects of

359 climate and weather combine to drive spatial and temporal variation in species abundance.

360 Model results showed that the direction of relationships between a variable and abundance

361 could vary depending on whether temporal or spatial variation was modelled separately

362 (Plummer et al. 2015). It also provided more detailed information about how annual

363 fluctuations in weather affected the spatial variation in population trends of five exemplar

364 bird species in Britain, and therefore may be used to infer additional responses to climate

365 change.

366   Responses to increasing winter and breeding season temperature in Britain have

367 generally been found to be positive, at least amongst resident species (Greenwood & Baillie

368   1991, Pearce-Higgins et al. 2015). These may be operating over multiple time-frames: from

369   immediate responses to direct extreme events, impacts of prey availability operating over

370   timescales of weeks to years, to long-term impacts of habitat change. Spatial variation in

371   climate may affect the long-term equilibrium of species abundances by impacting the average

372   habitat, food resources and indirectly the biotic competition. Temporal variation in weather

373   may be more likely to have immediate effects, such as changing survival or productivity,

374   through physiology, food availability, or breeding conditions (e.g. Robinson, Baillie & Crick

375   2007, Pearce-Higgins et al. 2015). Our proposed methods have the potential to assist in

376   disentangling these multiple processes in a single analysis, as illustrated by our responses for

377   five exemplar bird species, which we consider here in two groups.

378   Populations of the resident species, goldcrest, linnet and song thrush, would generally

379   be expected to have increased during our study period due to warmer temperatures leading to

380   increases in overwinter survival and reproductive success (Eglington & Pearce-Higgins 2012,

381   Pearce-Higgins et al. 2015). As expected, positive effects of preceding winter temperature

382   (W) were generally apparent for goldcrest and linnet across most of the linear models (Figure

383   4). However, in *mSTS_lBPWd* goldcrest and song thrush both showed large negative effects

384   of W_{space}, highlighting the potential divergence of spatial and temporal responses to

385   temperature. Linnet showed a large positive effect of W_{space} suggesting lower abundance in

386   locations with colder average winters (Swann et al. 2014).

387   For the migratory species considered, cuckoo and willow warbler, models predicted

388   strong population declines in the south and increases in the north (Figure 6) which concur

389   with previous studies (Ockendon et al. 2012; Morrison et al. 2013). Variation in these

390   populations may be partly influenced by conditions on migration (Hewson et al. 2016) or on

391   their African wintering grounds (Johnston et al. 2016). However, we found strong spatial

392   trends in abundance which suggests there was likely also a strong impact of breeding season

393   environmental variables for these species, including summer temperature, which is

394   quadratically associated with spatial variation in willow warbler productivity (Eglington et al.

395   2015), and rainfall. Positive effects of temporal variation in breeding season temperature B_{time}

396   and negative effects of spatial variation B_{space} – as shown for willow warbler – add weight to

397   previous suggestions for poleward shift in distribution and negative impacts of warming on

398   the breeding grounds in the south. This may be affecting at least some long-distance migrant

399   species, potentially through more lagged effects (e.g. Pearce-Higgins et al. 2015).

400   Overall, the spatial terms indicated larger effects than the temporal and residual terms

401   in the models for most species and effects. If this was due to the generally wider range of

402   values for the spatial term compared to the corresponding temporal or residual terms (see x-

403   axes in Fig. 5), we would expect the pattern of, e.g., the smooth for the temporal component

404   to resemble the smooth for the spatial component for the equivalent range of values. This was

405   not generally the case (see Appendix 1, Figure A1.2). Given that spatial variation in

406   environmental variables are here found to be most important, this may suggest limited

407   evidence for local adaptation in these populations and environmental covariates. It does

408   suggest that spatial variation in environment cannot be reliably used as a proxy for temporal

409   variation in environment. This finding promotes further caution in predicting species

410   distribution and abundance in future climate, based on relationships that are largely

411   determined by environmental variation over space. In some cases the direction or shape of

412   effect differed across the decomposed variables, suggesting that different ecological

413   processes govern spatial and temporal patterns in abundance.

414   The potential user of our methods should consider if the quasi-Poisson is the

415   appropriate error structure for fitting models to their data. Alternatives are the Poisson (if data

416   are not over/underdispersed) or the negative binomial. The advantage of the Poisson or

417   negative binomial over the quasi-Poisson is that they are standard distributions and allow

418 using, e.g., AIC for model selection. For the BBS data, we chose quasi-Poisson as it

419 accommodates overdispersion and, in contrast to the negative binomial, the underdispersion

420 encountered for cuckoo (and several other species in the BBS data not presented here). It also

421 provided a better fit to the data compared to the negative binomial (as evidenced by smaller

422 mean-squared errors and lower cross-validation scores; Appendix 3).

423   A remaining potential issue for fitting complex models like ours is overfitting, i.e. that

424 unnecessarily complex models might appear to be preferred over simpler models regardless

425 of whether the additional level of complexity reflects the underlying truth. Where this issue

426 appears to be a substantive concern, it should be addressed in the usual way via testing

427 existing model selection tools in simulation studies and, where necessary, trialling potential

428 new model selection tools.

429   Given the multi-dimensional nature of climate change, developing tools for

430 incorporating multiple climatic factors will ultimately be required to more fully model the

431 overall impact of climate change on species' populations. The general increase in well-

432 designed species recording schemes will provide a greater range of response variables; the

433 expansion of land-based and aerial earth observation has potential to provide covariate data

434 with fine temporal and spatial resolution. Hence, it is reasonable to expect these methods to

435 become increasingly applicable in future. Our methods can be extended in many ways, e.g.

436 by introducing some spatial averaging into the covariates since it is unlikely that only the

437 values at a location affect the species there. Hence, whilst the methods introduced in this

438 paper are a substantial advance, there is further development to be undertaken to extract

439 further information from spatio-temporal species-environment data sets.

440

441 **References**

442   Araújo, M.B., Pearson, R.G., Thuiller, W. & Erhard, M. (2005) Validation of species–climate

443   impact models under climate change. *Global Change Biology* 11, 1504-1513.

444   Beale, C.M., Baker, N.E., Brewer, M.J. & Lennon, J.J. (2013) Protected area networks and

445   savannah bird biodiversity in the face of climate change and land degradation.

446   *Ecology Letters,* 16(8), 1061-1068.

447   Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W. & Courchamp, F. (2012) Impacts of

448   climate change on the future of biodiversity. *Ecology Letters,* 15, 365-377.

449   Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. & Thomas, L.

450   (2001) *Introduction to distance sampling*. Chapman & Hall, London.

451   Buckland, S.T., Baillie, S.R., Dick, J.M., Elston, D.A., Magurran, A.E., Scott, E.M., Smith,

452   R.I., Somerfield, P.J., Studeny, A.C. & Watt, A. (2012) How should regional

453   biodiversity be monitored? *Environmental and Ecological Statistics,* 19, 601-626.

454   Chen, I.-C., Hill, J.K., Ohlemüller, R., Roy, D.B. & Thomas, C.D. (2011) Rapid range shifts

455   of species associated with high levels of climate warming. *Science*, **333**, 1024-1026.

456   Davey, C.M., Chamberlain, D.E., Newson, S.E., Noble, D.G. & Johnston, A. (2012). Rise of

457   the generalists: evidence for climate driven homogenization in avian communities.

458   *Global Ecology and Biogeography*, 21(5), 568-578.

459   Devictor, V., van Swaay, C., Brereton, T., Chamberlain, D., Heliölä, J., Herrando, S.,

460   Julliard, R., Kuussaari, M., Lindström, Å., Roy, D.B. & Schweiger, O. (2012).

461   Differences in the climatic debts of birds and butterflies at a continental scale. *Nature*

462   *Climate Change*, 2(2), 121-124.

463   Eglington, S.M. & Pearce-Higgins, J.W. (2012) Disentangling the relative importance of

464   changes in climate and land-use intensity in driving recent bird population trends.

465   *Plos One,* 7, e30407.

466 Eglington, S.M., Julliard, R., Gargallo, G., van der Jeugd, H.P., Pearce-Higgins, J.W., Baillie,

467   S.R. & Robinson, R.A. (2015) Latitudinal gradients in the productivity of European

468   migrant warblers have not shifted northwards during a period of climate change.

469   *Global Ecology & Biogeography,* 24, 427-436.

470 Greenwood, J.J. & Baillie, S.R. (1991). Effects of density-dependence and weather on

471   population changes of English passerines using a non-experimental paradigm. *Ibis*,

472   133, 121-133.

473 Gregory, R.D. & Baillie, S.R. (1998) Large-Scale Habitat Use of Some Declining British

474   Birds. *Journal of Applied Ecology*, 35, 785-799.

475 Harrison, P.J., Buckland, S.T., Yuan, Y., Elston, D.A., Brewer, M.J., Johnston, A. and

476   Pearce-Higgins, J.W. (2014) Assessing trends in biodiversity over space and time

477   using the example of British breeding birds. *Journal of Applied Ecology,* 51, 1650-

478   1660.

479 Harrison, P.J., Yuan, Y., Buckland, S.T., Oedekoven, C.S., Elston, D., Brewer, M.J.,

480   Johnston, A. and Pearce-Higgins, J.W. (2016) Quantifying turnover in biodiversity of

481   British breeding birds. *Journal of Applied Ecology*, 53, 469-478.

482 Hewson, C.M., Thorup, K., Pearce-Higgins, J.W. & Atkinson, P.W. (2016) Population

483   decline is linked to migration route in the common cuckoo. *Nature Communications,*

484   7:12296.

485 Johnston, A., Ausden, M., Dodd, A.M., Bradbury, R.B., Chamberlain, D.E., Jiguet, F.,

486   Thomas, C.D., Cook, A.S.C.P., Newson, S.E., Ockendon, N., Rehfisch, M.M., Roos,

487   S., Thaxter, C.B., Brown, A., Crick, H.Q.P., Douse, A., McCall, R.A., Pontier, H.,

488   Stroud, D.A., Cadiou, B., Crowe, O., Deceuninck, B., Hornman, M. & Pearce-

489   Higgins, J.W. (2013) Observed and predicted effects of climate change on species

490   abundance in protected areas, *Nature Climate Change*, 3(12), 1055-1061.

491   Johnston, A., Robinson, R.A., Gargallo, G. Julliard, R., van der Jeugd, H. & Baillie S.R.

492   (2016) Survival of Afro-Palaearctic passerine migrants in western Europe and the

493   impacts of seasonal weather variables. *Ibis* 158(3): 465-480.

494   Morrison, C.A., Robinson, R.A., Clark, J.A., Risely, K. & Gill, J.A. (2013) Recent

495   population declines in Afro-Palaearctic migratory birds: the influence of breeding and

496   non-breeding seasons. *Diversity and Distributions*, 19, 1051-1058.

497   Morton, D., Rowland, C., Wood, C., Meek, L., Marston, C., Smith, G., Wadsworth, R.,

498   Simpson, I.C., (2011) Final Report for LCM2007 - the new UK Land Cover Map.

499   Countryside Survey Technical Report No. 11/07 NERC/Centre for Ecology &

500   Hydrology 112pp. (CEH Project Number: C03259).

501   Ockendon, N., Hewson, C.M., Johnston, A. & Atkinson, P.W. 2012. Declines in British-

502   breeding populations of Afro-Palaearctic migrant birds are linked to bioclimatic

503   wintering zone in Africa, possibly via constraints on arrival time advancement. *Bird*

504   *Study,* 59, 111-125.

505   Ockendon, N., Johnston, A. & Baillie, S.R. 2014. Rainfall on wintering grounds affects

506   population change in many species of Afro-Palaearctic migrants. *Journal of*

507   *Ornithology,* 155, 905-917.

508   Pearce-Higgins, J.W., Eglington, S.M, Martay, B. & Chamberlain, D.E. (2015) Drivers of

509   climate change impacts on bird communities. *Journal of Animal Ecology*, 84, 943-54.

510   Plummer, K.E., Siriwardena, G.M., Conway, G.J., Risely, K. & Toms, M.P. (2015) Is

511   supplementary feeding in gardens a driver of evolutionary change in a migratory bird

512   species? *Global Change Biology*, 21, 4353-4363.

513   R core development team (2011) R: A language and environment for statistical computing R

514   foundation for statistical computing. Vienna, Austria. http://www.R-project.org.

515   Renwick, A.R., Massimino, D., Newson, S.E., Chamberlain, D.E., Pearce-Higgins, J.W. &

516   Johnston, A. (2012) Modelling changes in species' abundance in response to

517   projected climate change. *Diversity and Distributions,* 18, 121-132.

518   Richards, S.A. (2008) Dealing with overdispersed count data in applied ecology. *Journal of*

519   *Applied Ecology*, 45, 218-227.

520   Risely, K., Massimino, D., Newson, S.E., Eaton, M.A., Musgrove, A.J., Noble, D.G., Procter,

521   D. & Baillie, S.R. (2013) *The Breeding Bird Survey 2012*. BTO Research Report 645.

522   British Trust for Ornithology, Thetford.

523   Robinson, R.A., Baillie, S.R., & Crick, H.Q. (2007). Weather-dependent survival:

524   implications of climate change for passerine population processes. *Ibis*, 149(2), 357-

525   364.

526   Schummer, M.L., Kaminski, R. Raedeke, A.H. & Graber, D.A. (2010) Weather-related

527   indices of autumn-winter Dabbling Duck abundance in Middle North America.

528   *Journal of Wildlife Management* 74(1), 94-101.

529   Swann, R.L., Dillon, I.A., Insley, H. & Mainwood, T. (2014) Movements of Linnets Linaria

530   cannabina in northern Scotland. *Ringing & Migration* 29(1), 19-28

531   Talley, S.M., Coley, P.D. & Kursar, T.A. (2002) The effects of weather on fungal abundance

532   and richness among 25 communities in the Intermountain West. *BMC Ecology* 2, 1-

533   11.

534   WallisDeVries, M.F., Baxter, W. & Van Vliet, A.J.H. (2011). Beyond climate envelopes:

535   effects of weather on regional population trends in butterflies. *Oecologia*, 167(2),

536   559–571.

537   Warren, R., VanDerWal, J., Price, J., Welbergen, J.A., Atkinson, I., Ramirez-Villegas, J.,

538   Osborn, T.J., Jarvis, A., Shoo, L.P., Williams, S.E. & Lowe, J. (2013) Quantifying the

539 benefit of early climate change mitigation in avoiding biodiversity loss. *Nature*

540 *Climate Change* 3, 678-682.

541 Wood, S.N. (2006) *Generalized additive models: an introduction with R*. Chapman & Hall,

542 London.

543 Yoccoz, N.G., Nichols, J.D. & Boulinier, T. (2001) Monitoring of biological diversity in

544 space and time. *Trends in Ecology & Evolution,* 16, 446-453.

545

552

553 Data accessibility

554 The weather data used in this manuscript can be obtained from the Met Office website

555 http://www.metoffice.gov.uk/climate/uk/datasets/. The bird data used as an example in this

556 manuscript are available through the BTO's standard data request procedure

557 (http://www.bto.org/research-data-services/data-services/data-and-information-policy).

558

559 **Table 1.** For the analysed species, "L" indicates species for which late visit records were

560 used; "propM" gives the proportion of Britain included in the models; "prop0" gives the

561 proportion of zero counts in the data; "p.hat" gives the estimated detection probabilities and

562 $\hat{\theta}_{m1}$ and $\hat{\theta}_c$ give the estimated quasi-Poisson dispersion parameters for the baseline model

563 (*mSTS*, Table 2) and the most complex model used for calculating QAIC, respectively.

ID	species name	visit	propM	prop0	p.hat	$\hat{\theta}_{m1}$	$\hat{\theta}_c$
1	goldcrest		0.99	0.80	0.27	1.45	1.43
2	song thrush		0.99	0.48	0.50	2.01	1.99
3	linnet		0.97	0.73	0.31	3.58	3.54
4	cuckoo	L	0.94	0.96	0.76	0.31	1.00
5	willow warbler	L	1.00	0.67	0.44	2.48	2.46

564

Table 2. Summaries and δ-QAIC values of the models fitted to the counts of five species of breeding birds. Models are compared in three sections given in bold. Models beginning with '*mSTS*' contain the space-time smooth, models beginning with '*m_*' do not; '*l*' refers to linear effects of the weather covariates, e.g. '*lWBPd*' contains the linear effects of W and B and of the decomposed covariate P, i.e. P_{time}, P_{space}, P_{resid}; '*l*' is replaced with '*s*' for models containing smooth functions of the weather covariates. All models contain the land class and elevation covariates from eqn (1) (Section 2.2.1). The δ-QAIC values were obtained by subtracting the minimum QAIC value across all models for a species from each model's QAIC value. Smallest δ-QAIC for each section are in bold font.

Model	Weather covariates	Space-time smooth	Goldcrest	Song thrush	Linnet	Cuckoo	Willow warbler
Baseline model							
mSTS	none	Yes	652	304	529	44	409
A: Weather included as linear covariates							
mSTS_lW	linear W	Yes	642	300	491	44	394
mSTS_lB	linear B	Yes	645	298	395	45	411
mSTS_lP	linear P	Yes	653	295	437	**36**	411
mSTS_lWBP	linear W, B, P	Yes	608	293	364	38	388
mSTS_lWBPd	linear W, B, P_{space}, P_{time}, P_{resid}	Yes	574	290	336	41	**363**
mSTS_lWPBd	linear W, P, B_{space}, B_{time}, B_{resid}	Yes	403	264	320	41	384
mSTS_lwBPWd	linear B, P, W_{space}, W_{time}, W_{resid}	Yes	**206**	**224**	**199**	39	374
B: Weather included as smooth effects							
mSTS_sW	smooth of W	Yes	621	248	467	42	374
mSTS_sB	smooth of B	Yes	626	182	361	41	214
mSTS_sP	smooth of P	Yes	652	293	412	32	357
mSTS_sWBP	smooths of W, B, P	Yes	569	160	299	30	174

mSTS_sWBPd	smooths of *W*, *B*, P_{space}, P_{time}, P_{resid}	Yes	429	118	119	**0**	45
mSTS_sWPBd	smooths of *W*, *P*, B_{space}, B_{time}, B_{resid}	Yes	218	**0**	110	22	**0**
mSTS_sBPWd	smooths of *B*, *P*, W_{space}, W_{time}, W_{resid}	Yes	**0**	28	**0**	30	49
C: Models without the space-time smooth							
*m_*	none	No	3621	2189	2583	672	8763
m_sW	smooth of *W*	No	3448	2116	2502	650	8456
m_sB	smooth of *B*	No	3603	1924	2322	596	4681
m_sP	smooth of *P*	No	3592	2139	2433	546	8013
m_sWBP	smooths of *W*, *B*, *P*	No	3405	1795	2094	496	4375
m_sWBPd	smooths of *W*, *B*, P_{space}, P_{time}, P_{resid}	No	3095	**1622**	1871	339	3526
m_sWPBd	smooths of *W*, *P*, B_{space}, B_{time}, B_{resid}	No	3170	1656	1859	**311**	**3330**
m_sBPWd	smooths of *B*, *P*, W_{space}, W_{time}, W_{resid}	No	**2974**	1742	**1825**	441	4135

573

574

575

576

Figure 1. Weather variable *W* (centred monthly means of the daily minimum temperature during the preceding winter) decomposed into the temporal (top left), spatial (right) and residual variation components (bottom left panel). What is shown e.g. for 1994 entails the temperatures from December 1993 and January-February 1994.
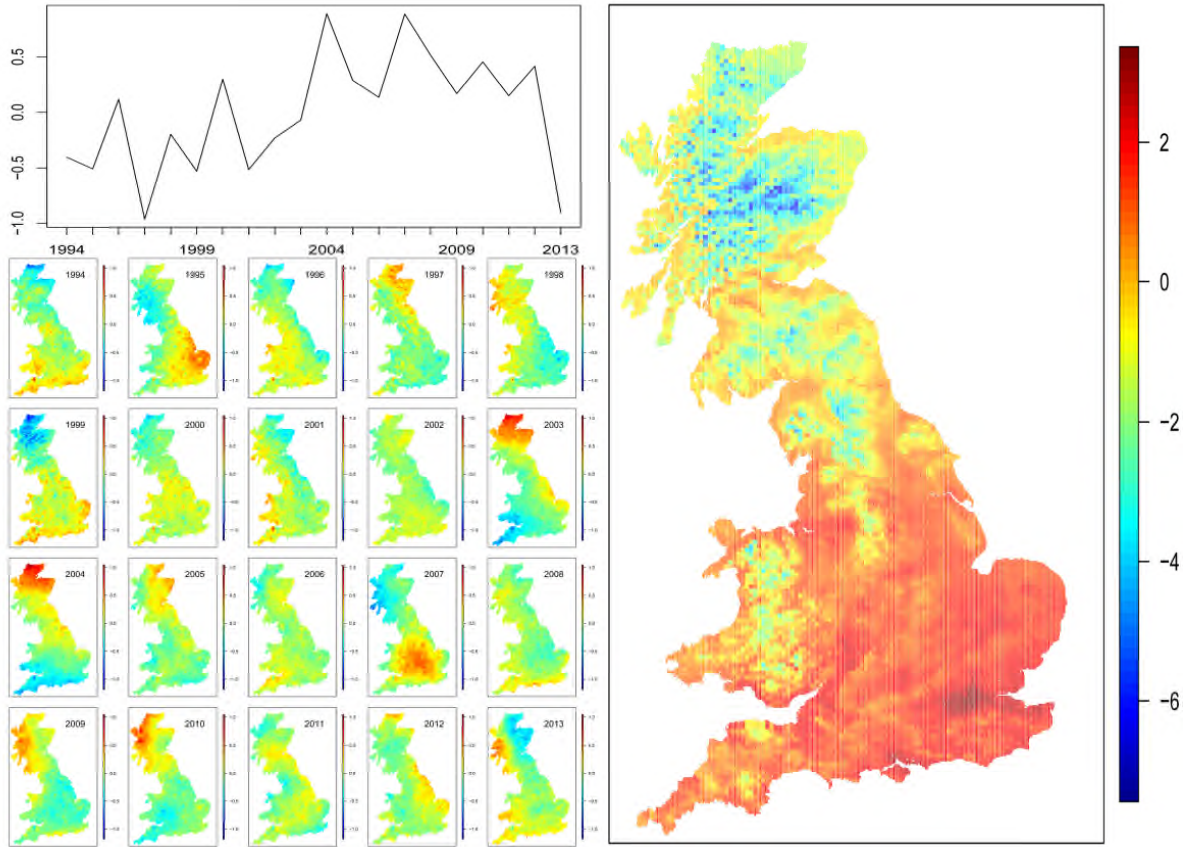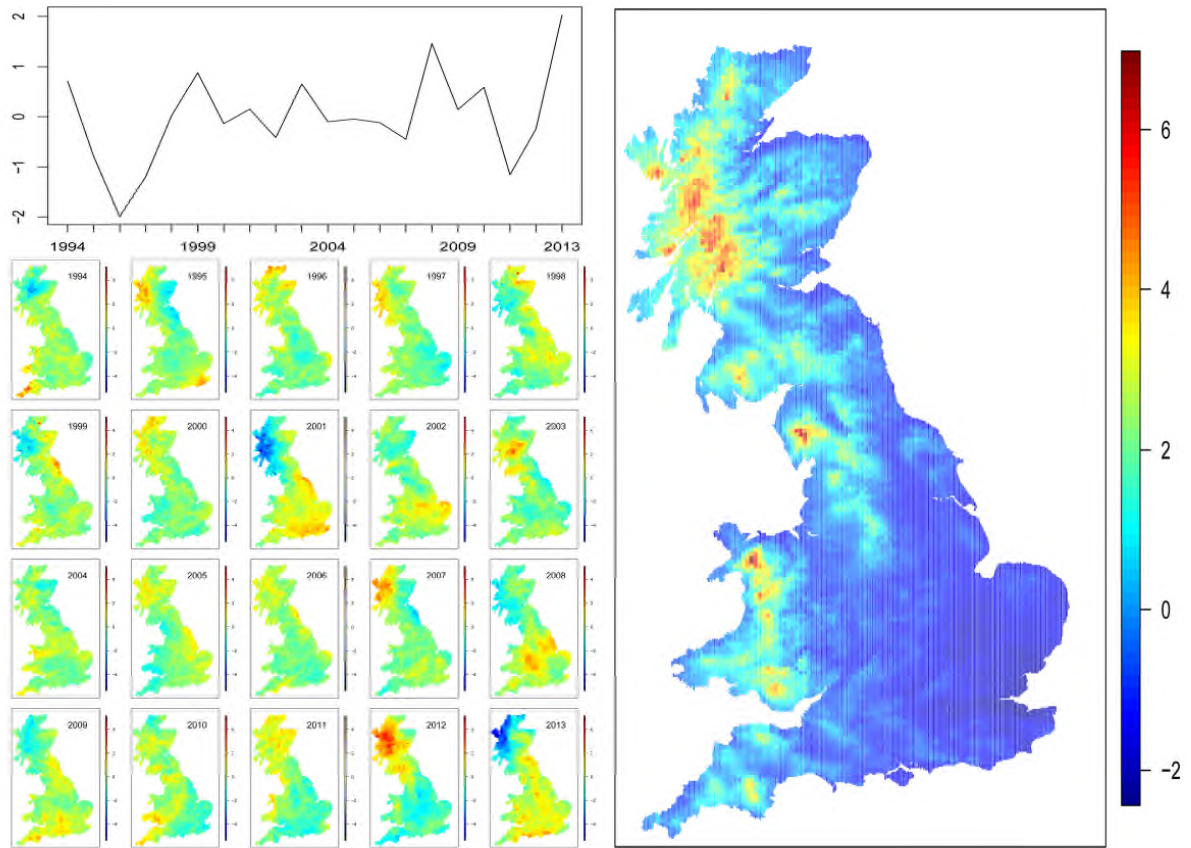
Figure 2. Weather variable *B* (centred monthly means of the daily mean temperature during the preceding breeding season) decomposed into the temporal (top left), spatial (right) and residual variation components (bottom left panel). What is shown e.g. for 1994 entails the temperatures April-July in 1993.

Figure 3. Weather variable *P* (centred square root of the monthly means of the total rainfall during the preceding breeding season) decomposed into the temporal (top left), spatial (right) and residual variation components (bottom left panel). What is shown e.g. for 1994 entails the rainfall from April-July in 1993.

Figure 4. Regression coefficients (and 95% confidence intervals) corresponding to linear effects of the weather covariates after standardisation (multiplying with the standard deviation of the covariate), presented as one panel for each species, with model names defined in Table 2. Dotted lines separate different models, dots give the estimates and horizontal lines represent the uncertainty of the estimates (the endpoints being the estimate±(SEx1.96)). Coefficients for *W* or its decomposed version (time top, space middle and anomaly bottom) are represented in green, for *B* in blue and for *P* in red. In the figure, estimates and interval endpoints are multiplied by the standard deviation of the respective covariate to allow the relative effects of the covariates to be assessed.

601   **Figure 5.** Comparison of smooths between centred and decomposed weather variables. For

602   each species, the smooths for the decomposed covariates from the best fitting model are

603   compared with the smooth of the corresponding centred covariate from model *mSTS_sWBP*.

604

605   **Figure 6.** Log-abundance estimates for the five bird species predicted for four selected years

606   using the best fitting model. Predictions were made only for areas included in the analysis

607   (see Section 2.1.1 and Table 1).

608

609

Figure 1. Weather variable *W* (centred monthly means of the daily minimum temperature during the preceding winter) decomposed into the temporal (top left), spatial (right) and residual variation components (bottom left panel). What is shown e.g. for 1994 entails the temperatures from December 1993 and January-February 1994.

610

611

612

613

614

615

616

Figure 2. Weather variable *B* (centred monthly means of the daily mean temperature during

the preceding breeding season) decomposed into the temporal (top left), spatial (right) and

residual variation components (bottom left panel). What is shown e.g. for 1994 entails the

temperatures April-July in 1993.

621

622

Figure 3. Weather variable *P* (centred square root of the monthly means of the total rainfall during the preceding breeding season) decomposed into the temporal (top left), spatial (right) and residual variation components (bottom left panel). What is shown e.g. for 1994 entails the rainfall from April-July in 1993.
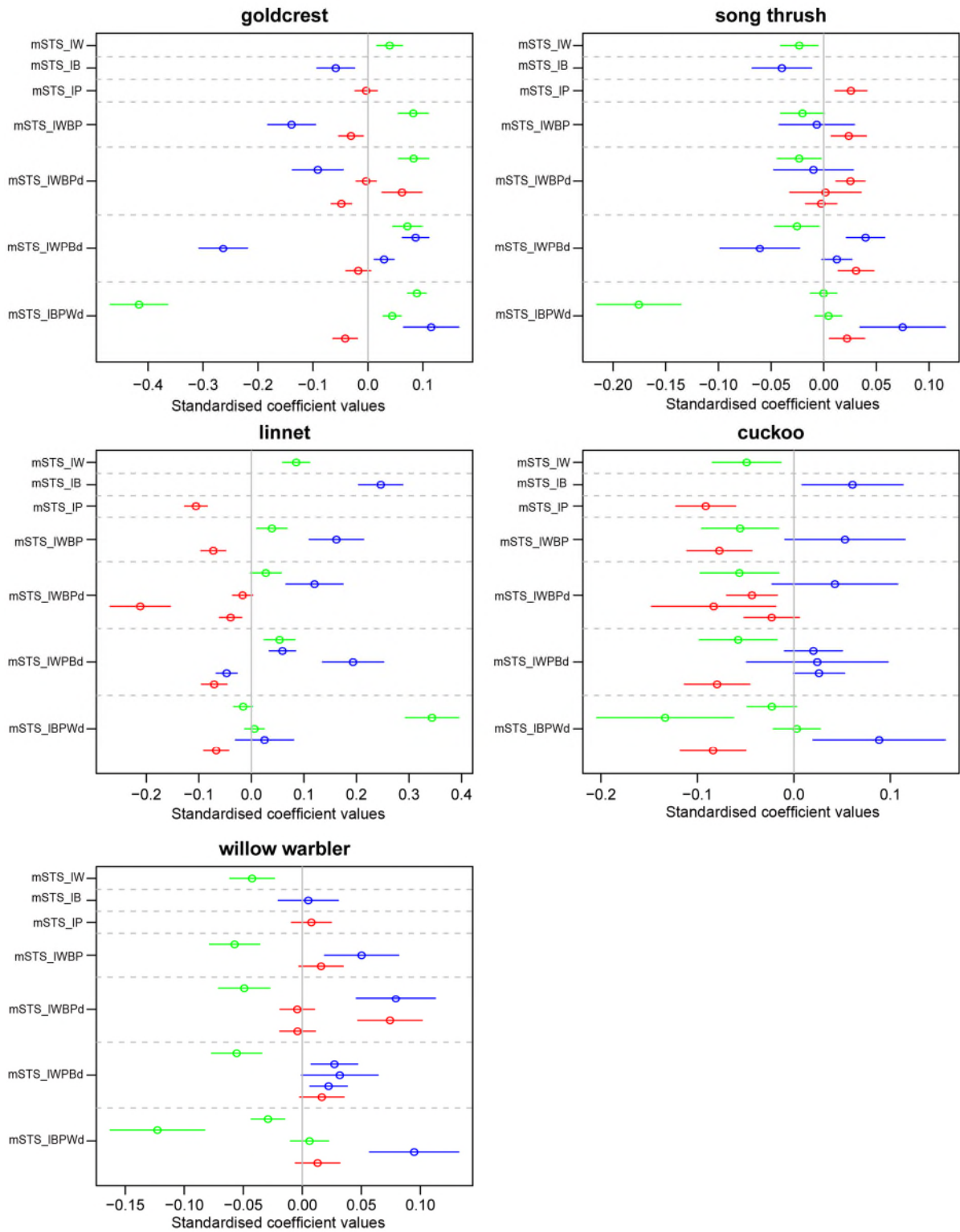
Figure 4. Regression coefficients (and 95% confidence intervals) corresponding to linear effects of the weather covariates after standardisation (multiplying with the standard deviation of the covariate), presented as one panel for each species, with model names defined in Table 2. Dotted lines separate different models, dots give the estimates and

635   horizontal lines represent the uncertainty of the estimates (the endpoints being the

636   estimate$\pm$(SEx1.96)). Coefficients for W or its decomposed version (time top, space middle

637   and anomaly bottom) are represented in green, for B in blue and for P in red. In the figure,

638   estimates and interval endpoints are multiplied by the standard deviation of the respective

639   covariate to allow the relative effects of the covariates to be assessed.

640
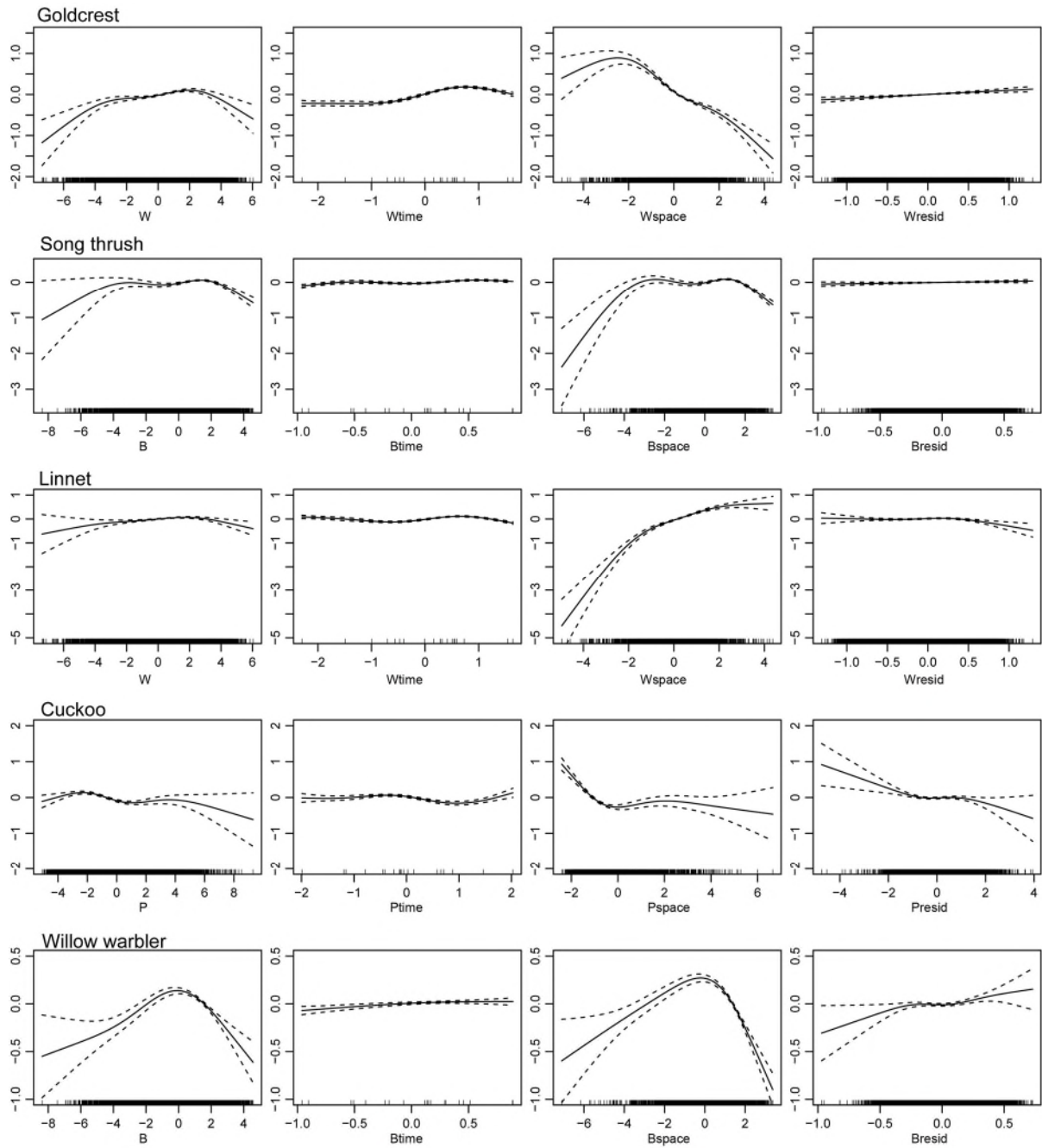
Figure 5. Comparison of smooths between centred and decomposed weather variables. For each species, the smooths for the decomposed covariates from the best fitting model are compared with the smooth of the corresponding centred covariate from model *mSTS_sWBP*.

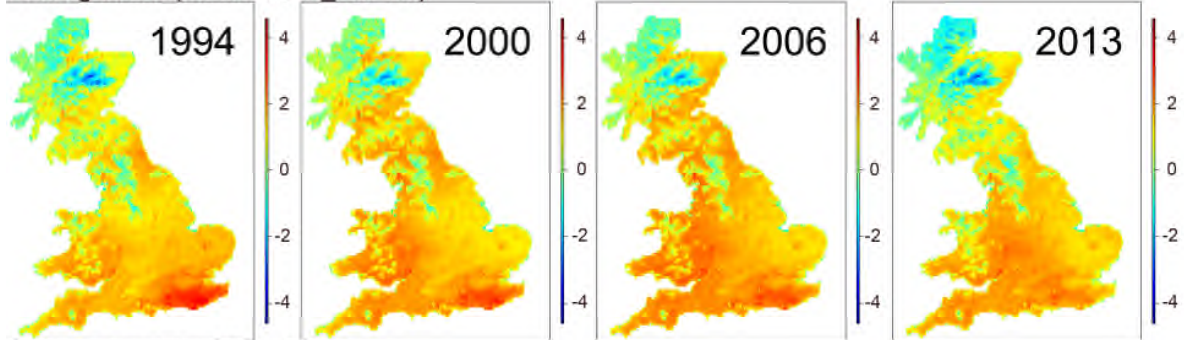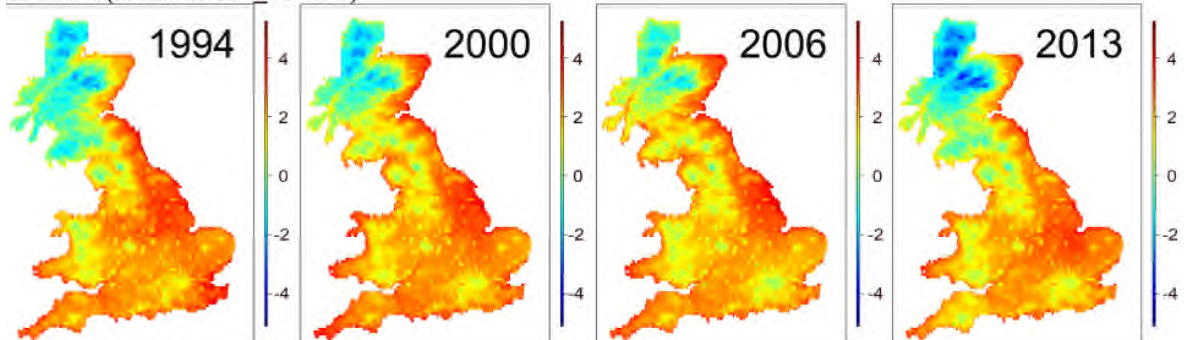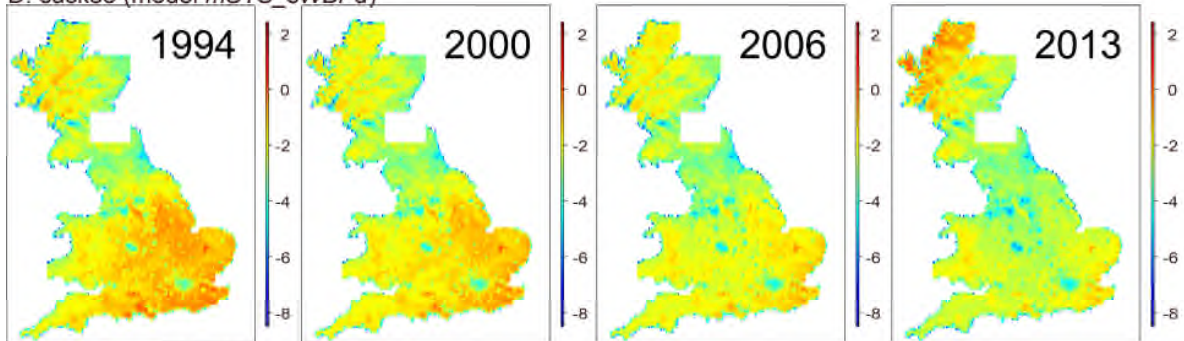A: goldcrest (model *mSTS_sBPWd*)
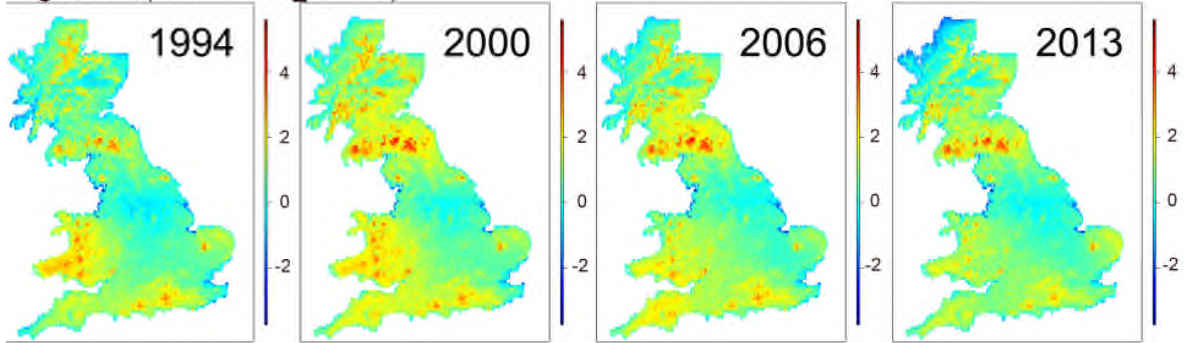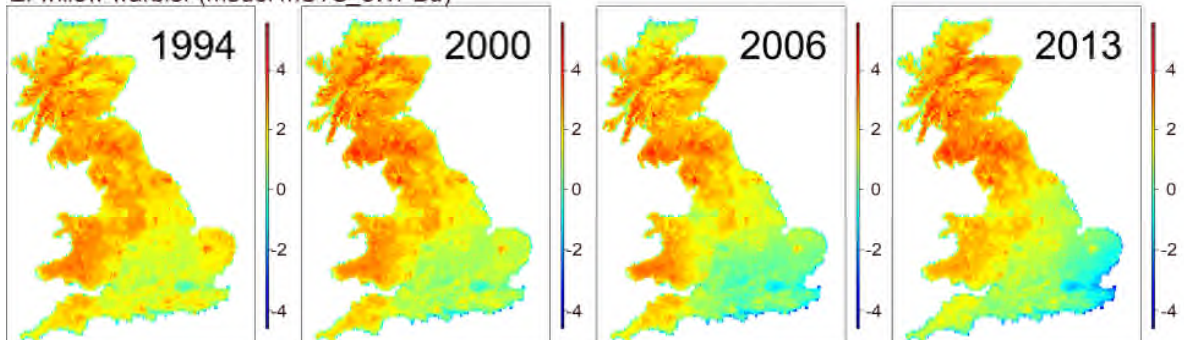
B: song thrush (model *mSTS_sWPBd*)

C: linnet (model *mSTS_sBPWd*)

D: cuckoo (model *mSTS_sWBPd*)

E: willow warbler (model *mSTS_sWPBd*)

648

649   **Figure 6.** Log-abundance estimates for the five bird species predicted for four selected years

650   using the best fitting model. Predictions were made only for areas included in the analysis

651   (see Section 2.1.1 and Table 1).

652

Authors' contribution statement

All authors contributed to developing the methodology of covariate decomposition; CSO lead the analyses and writing of the manuscript. Major parts of the R code were first developed by PJH. Species were selected by SF, DAE, MKB, AJ and SB. Data were provided by AJ and JWPH. All authors contributed critically to the drafts and gave final approval for publication.