Identifying environmental drivers of insect phenology across space and time: *Culicoides* in Scotland as a case study

K.R. Searle[1]*, A. Blackwell[2], D. Falconer[2], M. Sullivan[2], A. Butler[3] and B.V. Purse[1]

[1]Centre for Ecology and Hydrology, Bush Estate, Edinburgh, EH26 0QB: [2]Advanced Pest Solutions Ltd, Prospect Business Centre, Gemini Crescent, Dundee Technology Park, Dundee, DD2 1T: [3]Biomathematics and Statistics Scotland, James Clerk Maxwell Building, The King's Buildings, Mayfield Road, Edinburgh, EH9 3JZ

Abstract

Interpreting spatial patterns in the abundance of species over time is a fundamental cornerstone of ecological research. For many species, this type of analysis is hampered by datasets that contain a large proportion of zeros, and data that are overdispersed and spatially autocorrelated. This is particularly true for insects, for which abundance data can fluctuate from zero to many thousands in the space of weeks. Increasingly, an understanding of the ways in which environmental variation drives spatial and temporal patterns in the distribution, abundance and phenology of insects is required for management of pests and vector-borne diseases. In this study, we combine the use of smoothing techniques and generalised linear mixed models to relate environmental drivers to key phenological patterns of two species of biting midges, *Culicoides pulicaris* and *C. impunctatus*, of which *C. pulicaris* has been implicated in transmission of bluetongue in Europe. In so doing, we demonstrate analytical tools for linking the phenology of species with key environmental drivers, despite using a relatively small dataset containing over-dispersed and zero-inflated data. We demonstrate the importance of landcover and climatic variables in determining the seasonal abundance of these two vector species, and highlight the need for more empirical data on the effects of temperature and precipitation on the life history traits of palearctic *Culicoides* spp. in Europe.

Keywords: climate, insect phenology, landcover, spatio-temporal patterns, vector-borne disease

(Accepted 28 May 2012; First published online 30 July 2012)

Introduction

The way in which the distribution and phenology of organisms varies over space and time is of fundamental significance to many questions in ecology. Spatial heterogeneity in resources and environmental conditions drive geographical patterns in the ecology of all species, determining variation in community and population processes. Understanding how spatial variation in environmental conditions affects the demography and population dynamics of species is critical to effective management and conservation. This is particularly important in the face of rapidly changing environments and climate change (Vinatier *et al.*, 2011).

In general, understanding how environmental variation drives species phenology relies upon matching patterns of the

---

*Author for correspondence
Fax: +44 (0)131 4453943
E-mail: katrle@ceh.ac.uk

abundance of a species over space and time with underlying environmental correlates. However, this approach involves inherent analytical difficulties associated with spatial and temporal autocorrelation (Hoeting, 2009). Moreover, for highly abundant and temporally fluctuating species such as insects, problems arise from dealing with overdispersed and skewed data, and data containing a large proportion of zeros (e.g. Sileshi, 2006). Understanding the environmental drivers of phenology and abundance in many insect species is increasingly important for pest management (Rodeghiero & Battisti, 2000; Eber, 2004) and assessing risk of vector-borne diseases (Gern *et al.*, 2008; Takken *et al.*, 2008; Yang *et al.*, 2008; Parham & Michael, 2010). Vector abundance and seasonal dynamics are important determinants of variation in the risk of vector-borne infections, forming key components of R0 models for assessing disease spread (Heffernan *et al.*, 2005). For orbiviral pathogens of livestock, the periods of highest disease impact each year and the ability to persist between years is intimately linked to the climate-mediated seasonal dynamics of their biting midge vectors, *Culicoides* spp. (Diptera: Ceratopogonidae). This is particularly true in temperate systems where there is typically a vector-free period of several months over winter.

Methodologies for analysing insect phenology data across space and time include the use of generalised additive models (e.g. Saldaña *et al.*, 2007; Hodgson *et al.*, 2010) or population models (e.g. Yang *et al.*, 2008, 2009) to examine the influence of environmental drivers on abundance patterns of insect species. However, these methods typically require long time-series of data, and the preponderance of such studies has focused on tropical species that reproduce year round. Therefore, disentangling the importance of environmental drivers in different critical periods of the year for seasonal species remains elusive. Previously, methods for analysing the seasonality and survival of blood-sucking insects in general, and *Culicoides* in particular, have relied upon parity-based methods (MacDonald, 1957; Birley & Boorman, 1982; Holmes & Birley, 1987; Lord & Baylis, 1999). Parity-based methods use sampling of parous and nulliparous individuals within a population to estimate the survival rate of haematophagous insects. However, these methods have rarely been applied to understanding the role of environmental variables in driving patterns of seasonality and survival. Parity-based methods have provided important advances in the understanding of *Culicoides* phenology, but are not easily extended to analyses of phenological variation over space and time, and are bound by several important limiting assumptions: daily sampling data is required or samples must be taken over an entire season, a constant proportion of the population must be sampled at each time step, there must be no bias in sampling between nulliparous and parous individuals, all losses from the population are assumed due to mortality, and the survival rate must be constant over the time period considered (Birley & Boorman 1982; Lord & Baylis, 1999). Moreover, recent work has questioned the use of pigmentation of the abdomen as a method for determining parity in *Culicoides* sp. (Braverman & Mumcuoglu, 2009). In light of these restrictions, and the growing urgency in achieving a better understanding of the environmental drivers of insect seasonality for disease risk and control, there is a need for new methodologies for examining trends in vector abundance across space and time that are less data-hungry, and less bound by restrictive assumptions.

Here, we present a methodology for linking environmental correlates with the phenology of two Palearctic species of *Culicoides*, using weekly abundance data collected across Scotland over a two-year period. *Culicoides pulicaris* is a farm-associated species and a candidate vector for recent Bluetongue epidemics in Europe (Caracappa *et al.*, 2003; Purse *et al.*, 2004). The geographic range of *C. pulicaris* extends across the Palaearctic region to Pakistan and the Middle East (Institute for Animal Health, 2011). *Culicoides impunctatus* is a heathland-associated species, occurring throughout the UK and northern Europe (Boorman, 1986), but extending less far south, into the Iberian Peninsula (Hill, 1947; Boorman, 1986). This species breeds in acidic, boggy ground and is infamous for causing high nuisance impact on the forestry and tourist industries (Hendry & Godwin, 1988), being responsible for 70–95% of *Culicoides* spp. attacks on humans (Boorman, 1986) across Scotland. We chose not to model the *C. obsoletus* complex in this study because of difficulties in interpretation associated with modelling a complex of species for which individual species-level data was not available. Both *C. pulicaris* and *C. impunctatus* exhibit large fluctuations in abundance on weekly timescales, and hence our methodology had to deal with highly overdispersed count data with many zero values. By implementing a combination of smoothing functions and Bayesian regression models, we aimed to identify critical timing of abundance peaks in each year of data collection for both species and to correlate these with environmental drivers. Most life history parameters of *Culicoides* are known to be sensitive to temperature and moisture levels (see Mellor *et al.*, 2000 for a review) though their precise impacts on Palearctic species are poorly described. Our framework incorporates freely available, remotely sensed correlates of temperature, previously found to be of value in explaining patterns in midge vector abundance (Baylis *et al.*, 2001; Tatem *et al.*, 2003; Purse *et al.*, 2004).

Methods

Trapping sites and methods

Scotland was divided into quadrats of 40×40 km, and 37 farms in separate quadrats were selected for weekly light-trap surveillance between April and November in 2008 and 2009. Geographical coordinates for all traps sites were determined using the GPS facility of a BlackBerry Storm 9500 Smartphone (Research in Motion Ltd, Egham, Surrey, UK). Catches were carried out using Onderstepoort-type blacklight traps with 8 W UV-light bulbs and down-draught suction motors. Traps were operated for one night per week, from approximately 19:00 to 08:00 h and were located outside close to livestock and positioned with the UV-light tube approximately 1.6 m above the ground. Insects were collected into a plastic beaker containing 200–300 ml of water with detergent as a wetting agent. *Culicoides* spp. were separated from other insects under a binocular dissecting microscope on the basis of wing patterns. *Culicoides pulicaris s.s.* and *C. impunctatus* were identified on the basis of wing patterns, after Campbell & Pelham-Clinton (1960) and Rawlings (1996), and then counted. For large catches (>300 individuals), a subsample of the catch was speciated as described in Van Ark & Meiswinkel (1992) and the total number of individuals multiplied by the inverse of this proportion to calculate an adjusted abundance. Only females were considered in the statistical analysis since male *Culicoides* do not blood feed and are unimportant either as vectors or as causes of biting nuisance and often make up a small proportion of adults caught in light traps ($\sim 5\%$ in this

Table 1. Environmental drivers used in analysis of *Culicoides* appearance and abundance models in Scotland during 2008 and 2009.

Metric	Source	Resolution	Calculation
Photoperiod			Day of the year when eight hours of daylight is passed
Accumulated degree days	MOD11A2 eight-day LST	1 km	Accumulated degree days above 10°C between November 1st and May 31st $$\sum_{\text{Nov 1st}}^{\text{May 31st}} [\text{mean temperature} - 10°C, 0]$$
Mean winter temperature	MOD11A2 eight-day LST	1 km	Mean temperature between November 1st and February 28th
Mean spring temperature	MOD11A2 eight-day LST	1 km	Mean temperature between March 21st and April 30th
Mean temperature during June and July	MOD11A2 eight-day LST	1 km	Mean temperature between June 1st and July 31st
Summed spring precipitation	ENSEMBLES	25 km	Summed precipitation during March, April and May
Summed precipitation during June and July	ENSEMBLES	25 km	Summed precipitation during June and July
Percent cover of agriculture	CORINE Landcover Map 2000 (land principally occupied by agriculture with significant areas of natural vegetation, class 21)	100 m	Percent cover in 1 km buffer zones around sites
Percent cover of moors and heathlands	CORINE Landcover Map 2000 (class 27)	100 m	Percent cover in 1 km buffer zones around sites
Percent cover of pasture	CORINE Landcover Map 2000 (class 18)	100 m	Percent cover in 1 km buffer zones around sites

study). Parous and nulliparous females were distinguished by examining the colouration of their abdominal walls as described by Dyce (1969). The abdominal wall tissue of gravid females develops burgundy-red colouration during the development of ovarian follicles. The colouration persists in the abdomen walls and can be used to distinguish nulliparous and parous *Culicoides* females.

Environmental drivers

We derived a series of temperature, precipitation and landcover metrics for each trapping site. Weekly mean temperatures for each site were calculated using the MODIS Land Surface Temperature and Emissivity eight-day product (MOD11A2) at a resolution of 1 km (Wan & Li, 1997). These data are distributed by the Land Processes Distributed Active Archive Center (LP DAAC), located at the US Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center (lpdaac.usgs.gov). Daily precipitation data was derived from the ENSEMBLES model at a resolution of 25 km (Haylock *et al.*, 2008). We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://eca.knmi.nl). Landcover estimates were taken from the CORINE 2000 Landcover Map (EEA, 2000). Using these time-series of temperature and precipitation data, we derived a suite of seasonal metrics for use in the analysis (table 1). Environmental covariates were centred prior to model fitting to improve the efficiency of sampling in the MCMC chain (McCarthy, 2007). This adjustment has no effect on the regression slope coefficients, although the intercept is altered, hence back-transformed posterior means for intercept estimates are reported here.

Models

Because of the difficulties inherent to analysing short time-series of abundance data from species as numerous and temporally variable as *Culicoides* midges, we adopted a two-phase modelling approach. We first fit generalised additive models (GAMs) to a three-week moving average of the number of nulliparous midges caught per week at each trapping location. These models assumed a Poisson distribution with a log link and used a three-week moving average of the raw count data to reduce spurious temporal changes in abundance due to variation in meteorological conditions during the weekly trapping effort. A separate model was fit for each site, and year of data was collected to identify the peaks in abundance for each trapping location in each year of data collection. These models were used to identify the timing of the peak in abundance for different generational peaks for both species at each site in each year. GAMs used a spline smoothing function and all models were fit in R using the 'mgcv' package (Wood, 2011). For all sites, the degree of smoothing was automatically selected using generalised cross validation. Sites for which there were missing data (more than two consecutive weeks of trapping lost per year) were dropped from subsequent analyses.

Having identified the timing of the first and second generational peaks for each species, we derived phenological traits to depict the seasonal change in *Culicoides* abundance across Scotland: the timing of the first and, if present, second generational peaks, the abundance associated with the first generational peak, the abundance associated with the second generational peak, and the difference in abundance between the first and second generational peaks (log[abundance first peak]/log[abundance second peak]). Abundance estimates for each generational peak were taken as the
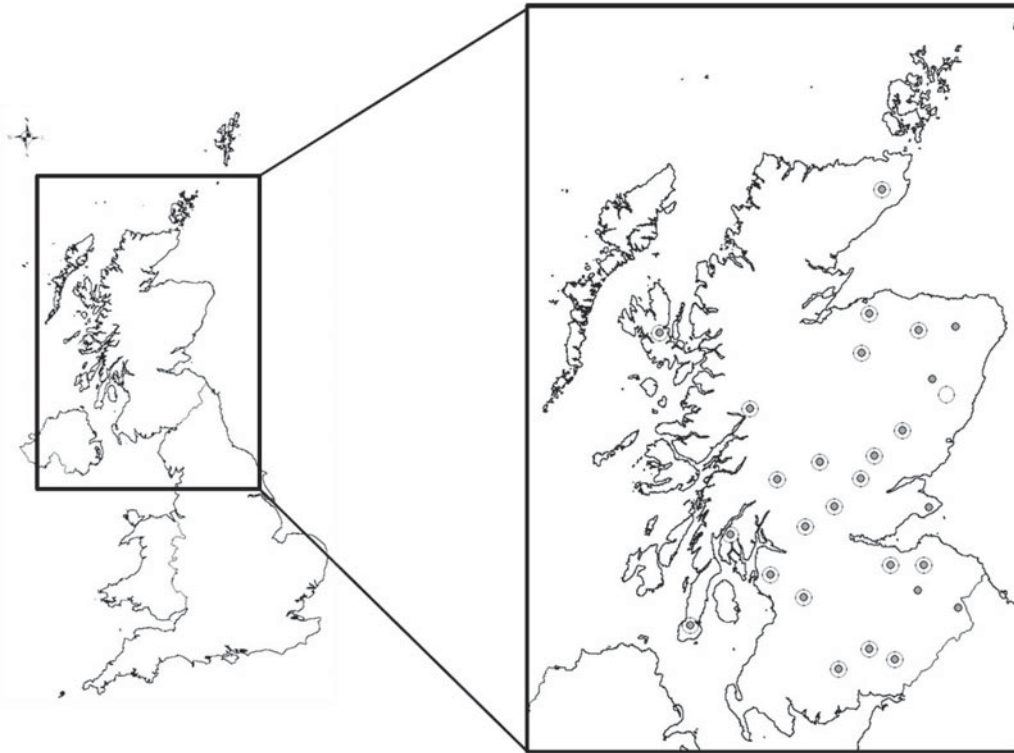
Fig. 1. Location of trapping sites for *C. pulicaris* (grey dots) and *C. impunctatus* (open circles).

six-week average in the total number of female midges trapped per site associated with the timing of each generational peak identified by the GAMs (none of the identified six-week periods for each generational peak overlapped for either species). We averaged counts of female midges over six weeks because this has previously been estimated as the generation period for *C. impunctatus* in Scotland (Kettle, 1950; Blackwell *et al.*, 1992).

The second phase of the analysis used linear mixed models (GLMMs) to relate these phenological traits to environmental drivers. The basic structure of these models related each phenological trait with a series of environmental drivers, assuming linear relationships. All models included a random intercept term for each year of measurement and a random site effect term to account for unstructured spatial heterogeneity not explicitly included in the regression model. Residuals from all models were checked for spatial autocorrelation (Moran's I, spdep package in R).

Correlation analyses (Pearson's r, R software program) were used to check for correlations between all of the environmental variables used in the GLMMs. No pair of variables showed a correlation greater than 0.5 (Pearson's correlation), and so all possible combinations of variables were used in the models. Deviance information criterion (DIC: Speigelhalter *et al.*, 2002) was used to select variables to find the most parsimonious model for each phenological trait. DIC is a generalisation of the Akaike information criterion (AIC) and is derived as the mean deviance adjusted for the estimated number of parameters in the model, compromising between model fit and complexity, and providing a measure of out-of-sample predictive error (Gelman & Hill, 2007). The mean deviance can be calculated from the posterior mean log

likelihood, i.e. the average of the log likelihood values calculated from the parameters in each sample from the posterior. Smaller values of DIC are indicative of a better-fitting model, and we used delta DIC values to compare support in the data for each model. When comparing two models, a delta DIC value of less than two indicates substantial support for both models, a value of four to seven indicates considerably less support for the model with the larger DIC, and a value of greater than ten indicates essentially no support for the model with the larger DIC (McCarthy, 2007).

Timing of the first and second peaks in abundance

Models for the timing of the first and, if present, second generational peaks assumed that the data were drawn from an underlying normal distribution. We assumed that the mean of this distribution had a linear relationship with environmental drivers:

$$Y_{sy} \sim \text{Normal}(\mu_{sy}, \sigma^2)$$
$$\mu_{sy} = a_y + \sum_i b_i x_{syi} + \varepsilon_s$$

where $Y_{s,y}$ are the week of appearance at site s in year y; μ_{sy} is the mean of the normal distribution (the process model estimate for timing of the first peak at each site s in each year y); x_{syi} is the value of environmental variable i in year y at site s; σ^2 is the residual variance in the data not explained by site, year or environmental variables; a_y is the random intercept term for each year of measurement; b_i are the regression
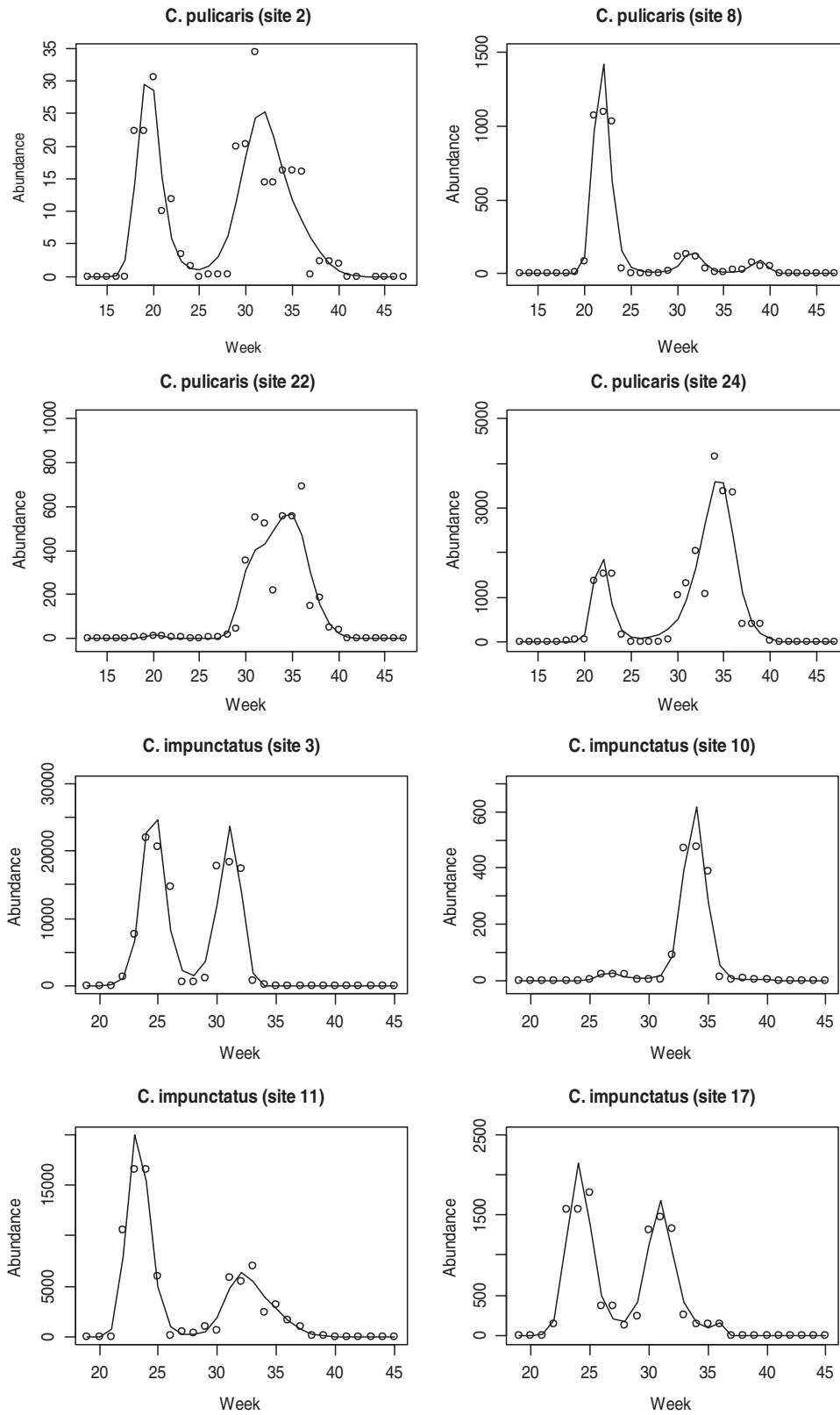
Fig. 2. Time-series for abundance counts (dots are three week moving average) and fitted Generalised Additive Models (solid line) for *C. pulicaris* and *C. impunctatus* at eight trapping sites during 2008.
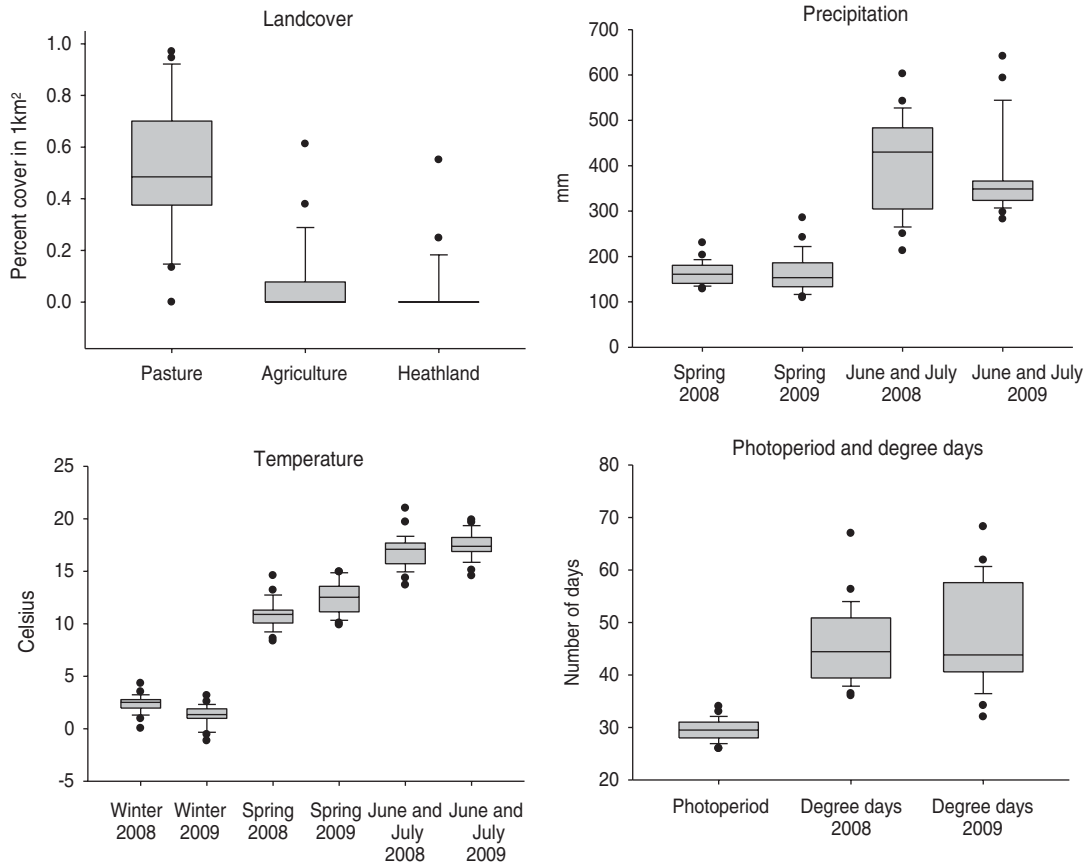
Fig. 3. Summary of environmental drivers used in GLMM models. Box plots show the median (central line), box denotes 25th and 75th percentiles, error bars represent 10th and 90th percentiles and dots are points outside the 10th and 90th percentiles.

coefficients for the environmental drivers; and ε_s accounts for unstructured spatial heterogeneity (random site effect).

Size of the first and second peaks

Counts from surveillance data are most appropriately modelled using a Poisson distribution. However, insect data are generally highly overdispersed, which presents difficulties in achieving convergence in Markov Chain Monte Carlo simulations when models include multiple random effects for spatial heterogeneity and overdispersion. Therefore, the abundance at each site per year during the first and second generational peaks was log transformed, and modelled assuming a normal distribution assuming a linear relationship with environmental drivers. The structure of the model for log abundance was identical to that for timing.

Difference in size between first and second peaks

The log difference in size between the abundance of the first and second generational peaks, $D_{s,y} = \log\left(\frac{\text{abundance of first peak}}{\text{abundance of second peak}}\right)$, was also assumed to have a normal distribution. The structure of the model was identical to that for timing. Larger values of $D_{s,y}$ indicate a larger first abundance relative to the second abundance, within a

particular site and year, suggesting greater mortality during the time interval between peak abundance sizes.

Inference

All models were fit using BRugs on the R platform (R-Development Core Team, 2009) and OpenBUGS software (Lunn *et al.*, 2009). Models were run for 100,000 iterations with three MCMC chains and a burn-in period of 50,000 iterations for each chain. MCMC samples were thinned to every 10th sample when necessary (due to autocorrelation in the MCMC chains), giving a total of 5000 (thinned) or 50,000 (unthinned) samples per chain. Model fit was checked using trace plots and the Gelmin-Rubin statistic for scale reduction factors (Gelmin & Rubin, 1992).

Mean parous rate

The mean parous rate, $\frac{\sum_{s=1}^{15} P_s}{\sum_{s=1}^{15} NP_s + P_s}$ (MacDonald, 1952), where P and NP are the total number of parous and nulliparous females trapped at each of the sites for *C. pulicaris* ($n=27$) and *C. impunctatus* ($n=23$), can be used to estimate survival rates when data has been collected over the entirety of the season (Birley & Boorman, 1982). We used the trapping
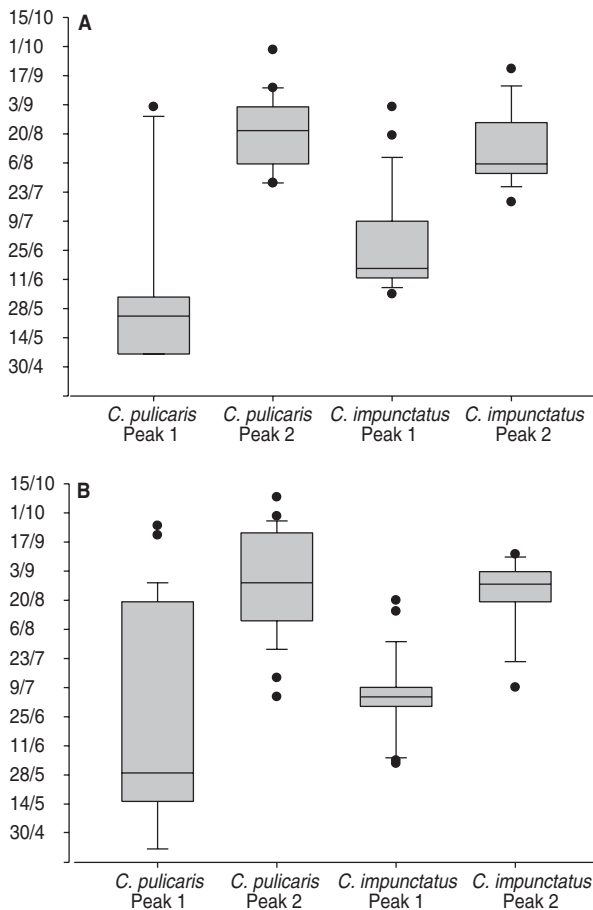
Fig. 4. Timing of first generational peak for *C. impunctatus* and *C. pulicaris* across Scotland in (A) 2008 and (B) 2009 identified by the fitted generalised additive models. Box plots show the median (central line), box denotes 25th and 75th percentiles, error bars represent 10th and 90th percentiles, and dots are points outside the 10th and 90th percentiles.

data to estimate mean survival rates between blood meals for *C. pulicaris* and *C. impunctatus*.

Results

Sites selected

After eliminating sites with gaps in time-series trapping data, 27 sites were selected for analysis for *C. pulicaris* and 23 for *C. impunctatus* in both 2008 and 2009 (fig. 1). Of the 27 sites used for *C. pulicaris*, 22 showed second generational peaks in 2008 and 21 in 2009; and, of the 23 sites used for *C. impunctatus*, 16 showed second generational peaks in 2008 and 12 in 2009. Model fit for the generalised linear models was good across all species and sites (adjusted $R^2 > 0.70$; fig. 2).

Culicoides pulicaris constituted approximately 16% in 2008 and 11% in 2009 of the total yearly catch of all females *Culicoides* across all 37 sites, whilst *C. impunctatus* made up approximately 25% in 2008 and 42% in 2009. The *C. obsoletus* group constituted approximately 5% in 2008 and 44% in 2009. Six other species were present in very low numbers: *C. punctatus*, *C. grisescens*, *C. deltus*, *C. fagineus*, *C. newsteadi* and *C. nubeculosus*.

Environmental variables

Of the three landcover types evaluated, pasture was the most prevalent around the trapping sites (fig. 3). Percent cover of agriculture and heathland was low, although a few sites had high percent cover values (fig. 3). Comparing annual mean environmental variables for the sites selected for analysis showed that sites accumulated similar amounts of degree days in 2008 and 2009; however, the variability (standard deviation) between sites was greater in 2009 than in 2008 (fig. 3). Mean temperature in winter was slightly higher in 2008 than in 2009, but mean temperature in the spring was slightly higher in 2009 than in 2008. Mean temperature in June and July was slightly lower, and more variable, in 2008 than in 2009 (fig. 3). Spring precipitation was similar in both years but was more variable in 2009 (fig. 3). Summed precipitation in June and July was greater in 2008 than in 2009 (fig. 3).

Timing of generational peaks

Averaged across all sites, *C. pulicaris* abundance peaked earlier than *C. impunctatus* abundance in both years (mean date of first generational peak: *C. pulicaris* 4th June 2008 and 17th June 2009; *C. impunctatus* 25th June 2008 and 1st July 2009), although the mean timing of the second generational peak was similar for both species (*C. pulicaris* 20th August 2008 and 26th August 2009; *C. impunctatus* 13th August 2008 and 26th August 2009) (fig. 4). In 2008, we detected a significant negative correlation between the timing of the first and second peaks in abundance (Pearson's $r = -0.48$, $t_{20} = -2.45$, $P = 0.023$); however, this correlation was not significant in 2009 (Pearson's $r = 0.38$, $t_{19} = 1.80$, $P = 0.087$). No significant correlations were detected between the timing of the first and second peaks for *C. impunctatus* (2008 Pearson's $r = 0.46$, $t_{14} = 1.91$, $P = 0.077$; 2009 Pearson's $r = 0.33$ $t_{10} = 1.10$, $P = 0.30$). We detected no significant spatial autocorrelation in the residuals from any of the best-fitting models.

There were five models selected with a delta DIC value of less than two (indicating approximately equal support for all models) for the timing of the first generational peak for *C. pulicaris* (table 2). These five models included photoperiod, percent cover of agriculture, spring temperature and percent cover of pasture. However, only photoperiod and percent cover of agriculture showed significant correlations (95% credible interval does not bridge zero) with the timing of the first generational peak (table 2). Sites that exceeded eight hours of daylight earlier in the year had earlier peaks, and sites with a greater percent cover of agriculture in the surrounding $1 \, km^2$ had earlier peaks (table 2).

Four models for the timing of the first generational peak for *C. impunctatus* had a delta DIC value of less than two (table 2). These models included photoperiod, degree days, winter temperature and percent cover of heathland (table 2). However, only photoperiod had a significant positive correlation with timing, indicating that sites at lower latitudes had earlier first generational peaks.

Three models for the timing of the second generational peak for *C. pulicaris* had delta DIC values of less than two, including combinations of June and July temperature, and percent cover of agriculture and pasture (table 3). Only mean June and July temperatures exhibited a significant positive

Table 2. Model results for timing of first peak in abundance for *C. pulicaris* and *C. impunctatus* in Scotland during 2008 and 2009. Posterior means for variable intercepts (*a*) for 2008 and 2009, and slope parameters for environmental variables (b_i) are shown with associated 95% credible intervals. The estimated number of parameters (*pD*), Deviance Information Criterion (DIC), and Gelmin-Rubin statistic for MCMC convergence (multivariate psrf) are shown for each model. Asterix (*) indicate significance for environmental variables. Δ_{DIC} refers to the difference in DIC values between each model and the model receiving the greatest support in the data. Dotted line separates model with approximately equal support ($\Delta_{DIC} < 2$) from those with considerably less support.

Species	Model	a 2008 a 2009	b_i	pD	DIC	Multivariate psrf	Δ_{DIC}
C. pulicaris	Photoperiod + agriculture	−28.06 (−54.35, −1.23) −26.51 (−52.57, 0.43)	1.78 (0.86, 2.67)* −16.13 (−28.87, −3.20)*	17.7	326.8	1.02	0
	Spring temperature	30.44 (18.03, 42.09) 33.00 (19.04, 46.01)	−0.68 (−1.73, 0.43)	23.2	327.7	1.03	0.9
	Photoperiod + agriculture + spring temperature	−27.40 (−58.61, 5.66) −25.77 (−58.37, 8.70)	1.77 (0.84, 2.67)* −16.17 (−29.60, −2.97)* −0.054 (−1.12, 1.14)	18.5	328.1	1.03	1.3
	Photoperiod	−20.19 (−48.51, 8.09) −18.68 (−47.06, 9.60)	1.47 (0.52, 2.43)*	19.2	382.2	1.02	1.4
	Photoperiod + agriculture + spring temperature + pasture	−22.94 (−62.70, 18.13) −21.23 (−61.98, 21.15)	1.67 (0.56, 2.75)* −16.15 (−29.75, −2.52)* −0.12 (−1.26, 1.11) −1.56 (−9.96, 6.78)	19.5	328.6	1.03	1.8
	Photoperiod + agriculture + spring temperature + pasture + winter temperature	−27.78 (−69.25, 16.77) −25.51 (−67.62, 19.27)	1.77 (0.60, 2.85)* −16.73 (−30.35, −2.60)* −0.063 (−1.22, 1.14) −1.39 (−9.69, 6.72) 0.61 (−1.32, 2.52)	19.3	331.3	1.03	4.5
	Photoperiod + degree days + winter temperature + spring temperature + agriculture + pasture	−29.47 (−70.95, 16.15) −26.89 (−68.94, 18.77)	1.79 (0.59, 2.88)* 0.081 (−0.13, 0.31) 0.62 (−1.27, 2.50) −0.31 (−1.66, 1.10) −16.96 (−30.19, −3.31)*	19.5	333.6	1.20	6.8
	Photoperiod + degree days + winter temperature + spring temperature + spring precipitation + agriculture + pasture	−21.44 (−70.94, 31.06) −18.95 (−68.83, 34.15)	1.66 (0.42, 2.85)* 0.067 (−0.16, 0.30) 0.47 (−1.49, 2.45) −0.36 (−1.72, 1.09) −0.018 (−0.078, 0.044) −17.56 (−30.80, −3.49)* −0.81 (−9.18, 7.32)	20.1	335.3	1.03	8.5
C. impunctatus	Photoperiod + degree days + winter temp	8.42 (−7.92, 25.61) 10.80 (−5.33, 27.65)	0.52 (0.029, 1.01)* −0.0083 (−0.13, 0.11) 1.24 (0.12, 2.35)	10.6	233.3	1.02	0
	Photoperiod	14.29 (−0.43, 29.11) 15.38 (0.67, 30.11)	0.41 (−0.092, 0.90)	11.0	234.2	1.01	0.9
	Photoperiod + degree days + winter temp + heath	8.06 (−8.51, 25.84) 10.38 (−6.17, 27.92)	0.55 (0.018, 1.06)* −0.0098 (−0.13, 0.11) 1.18 (−0.16, 2.46) −0.90 (−9.35, 7.29)	11.7	235.0	1.01	1.7
	Photoperiod + degree days	14.13 (−2.98, 32.59) 15.23 (−1.88, 33.73)	0.40 (−0.13, 0.92) 0.0054 (−0.13, 0.13)	11.6	235.3	1.02	2.0
	Photoperiod + degree days + winter temp + spring temp	8.85 (−9.72, 28.44) 11.31 (−7.49, 31.16)	0.52 (0.0077, 1.01)* −0.0018 (−0.15, 0.14) 1.24 (0.093, 2.34)* −0.055 (−0.90, 0.79)	11.1	235.4	1.02	2.1
	Photoperiod + degree days + spring temp	15.83 (−3.86, 37.34) 17.10 (−3.19, 39.17)	0.38 (−0.18, 0.92) 0.016 (−0.15, 0.18) 1.44 (0.11, 3.26)	11.8	237.0	1.01	4.0
	Photoperiod + degree days + winter temp + heath + spring temp	8.82 (−10.17, 29.25) 11.19 (−8.14, 31.93)	0.54 (−0.010, 1.07)* −0.0031 (−0.15, 0.14) 1.14 (−0.20, 2.44) −1.07 (−9.66, 7.24) −0.073 (−0.94, 0.78)	12.7	237.1	1.01	4.1
	Photoperiod + degree days + winter temp + heath + spring temp + spring ppt	10.90 (−16.49, 38.66) 13.27 (−14.51, 41.25)	0.50 (−0.16, 1.15) −0.0054 (−0.16, 0.15) 1.13 (−0.21, 2.47) −0.72 (−9.80, 8.08) −0.075 (−0.95, 0.80) −0.0036 (−0.041, 0.035)	13.3	238.9	1.01	5.6

relationship with the timing of the second peak, indicating that sites experiencing warmer summers had later second generational peaks.

For *C. impunctatus*, five models selected had delta DIC values of less than two, including June and July precipitation, June and July temperatures, and percent cover of heathland

Table 3. Model results for timing of second peak in abundance for *C. pulicaris* and *C. impunctatus* in Scotland during 2008 and 2009. Posterior means for variable intercepts (*a*) for 2008 and 2009, and slope parameters for environmental variables (b_i) are shown with associated 95% credible intervals. The estimated number of parameters (*pD*), Deviance Information Criterion (DIC), and Gelmin-Rubin statistic for MCMC convergence (multivariate psrf) are shown for each model. Asterix (*) indicate significance for environmental variables. Δ_{DIC} refers to the difference in DIC values between each model and the model receiving the greatest support in the data. Dotted line separates model with approximately equal support ($\Delta_{DIC} < 2$) from those with considerably less support.

Species	Model	*a* 2008 *a* 2009	b_i	*pD*	DIC	Multivariate psrf	Δ_{DIC}
C. pulicaris	June and July mean temperature	20.58 (8.59,32.63) 20.85 (8.07,33.69)	0.81 (0.096,1.54)*	7.8	221.1	1.03	0
	June and July mean temperature + agriculture	21.20 (8.61,33.91) 21.49 (8.19,34.91)	0.78 (0.031,1.53)* −1.66 (−8.44,5.19)	8.8	222.7	1.03	1.6
	June and July mean temperature + agriculture + pasture	21.30 (9.09,33.85) 21.75 (8.70,35.02)	0.69 (−0.067,1.44) −1.02 (−7.68,5.50) 2.53 (−1.38,6.46)	9.8	222.7	1.03	1.6
	June and July mean temperature + agriculture + pasture + June and July precipitation	24.80 (10.71,39.36) 25.04 (10.37,40.04)	0.65 (−0.11,1.39) −1.82 (−8.78,5.19) 3.59 (−0.74,8.10) −0.017 (−0.054,0.018)	11.0	223.2	1.03	2.1
C. impunctatus	June and July ppt	36.32 (30.31,42.37) 37.89 (32.10,43.66)	−0.019 (−0.050,0.011)	7.7	132.4	1.01	0
	June and July temp	28.26 (14.55,42.14) 29.74 (15.19,44.30)	0.27 (−0.57,1.09)	8.4	133.0	1.01	0.6
	June and July ppt + June and July temp	32.30 (17.10,47.33) 33.64 (17.94,49.25)	−0.019 (−0.050,0.012) 1.07 (0.077,2.53)	8.9	133.4	1.02	1.0
	heath	32.54 (31.10,33.97) 34.26 (32.57,35.93)	0.75 (−5.52,7.07)	8.1	133.8	1.01	1.4
	June and July ppt + heath	36.31 (29.95,42.72) 37.85 (31.65,44.03)	−0.019 (−0.051,0.013) −0.072 (−6.37,6.31)	8.6	134.1	1.02	1.7
	June and July temp + heath	28.18 (14.08,42.56) 29.64 (14.66,44.85)	0.27 (−0.61,1.13) 0.44 (−6.15,7.02)	9.4	134.7	1.01	2.3
	June and July ppt + June and July temp + heath	32.04 (16.50,47.68) 33.33 (17.20,49.54)	−0.019 (−0.052,0.014) 0.26 (−0.62,1.11) −0.27 (−6.81,6.31)	9.8	135.2	1.01	2.8

(table 3). However, none of these variables showed a significant correlation with timing of the second peak for this species.

Model fit for the timing of the first peak for *C. pulicaris* was good ($R^2 = 0.91$, $F_{2,24} = 125$, $P < 0.001$) and was reasonable for *C. impunctatus* ($R^2 = 0.61$, $F_{2,20} = 18.0$, $P < 0.001$). Model fit for the timing of the second peak was poor for *C. pulicaris* ($R^2 = 0.28$, $F_{2,17} = 4.7$, $P = 0.023$) but good for *C. impunctatus* ($R^2 = 0.80$, $F_{2,6} = 16.8$, $P = 0.0035$).

Size of generational peaks

Culicoides impunctatus produced consistently larger peaks in abundance compared to *C. pulicaris* (mean size of first peak: *C. impunctatus*, 5284 in 2008 and 7603 in 2009; *C. pulicaris*, 1390 in 2008 and 748 in 2009; mean size of second peak: *C. impunctatus*, 2417 in 2008 and 3088 in 2009; *C. pulicaris*, 369 in 2008 and 1380 in 2009) (fig. 5).

No variables showed a significant correlation with the size of the first generational peak for *C. pulicaris* (table 4). The two best models included combinations of degree days, spring precipitation, pasture and spring temperature. Two models for

C. impunctatus had delta DIC values of less than two and included combinations of photoperiod, degree days, spring precipitation and percent cover of heathland (table 4). Only photoperiod showed a significant positive correlation with the size of the first peak for this species, suggesting that sites which exceeded eight hours of daylight earlier in the year had smaller abundances during the first generational peak.

It was not possible to select between the four environmental variables used in models for the size of the second generational peak for *C. pulicaris*; all combinations of variables had DIC values within two of each other (table 5). However, the percent cover of pasture showed a significant positive relationship with the size of the second peak for this species, indicating that sites with a greater cover of pasture in the surrounding 1 km² had higher abundances during the first generational peak. We found a significant positive relationship between the size of the first and second peaks in abundance for *C. pulicaris* that persisted in both years of study (2008 Pearson's r = 0.52, $t_{20} = 2.71$, $P = 0.013$; 2009 Pearson's r = 0.45, $t_{19} = 2.17$, $P = 0.043$). We, therefore, included the size of the first peak in the best-fitting model identified during environmental variable selection; and, indeed, this
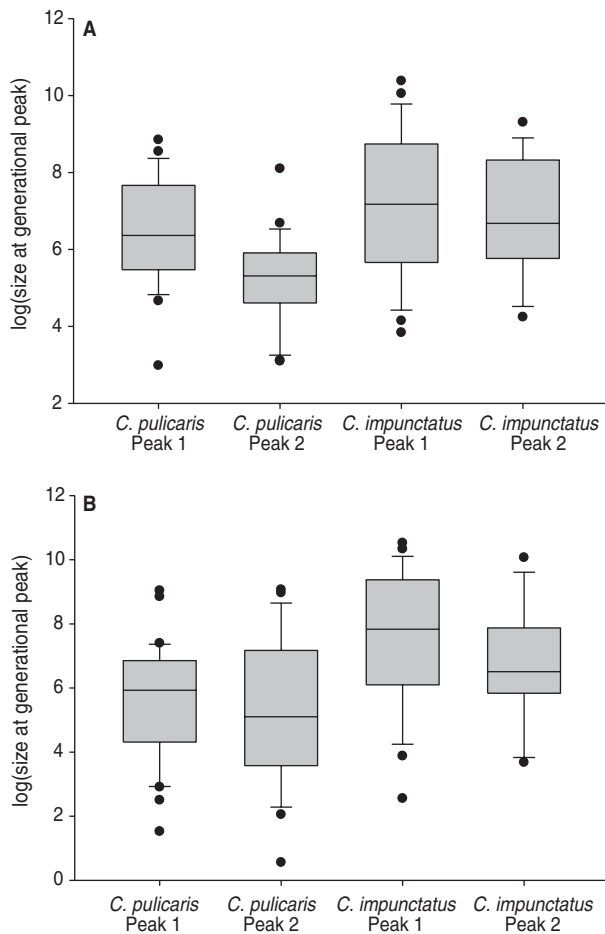
Fig. 5. Mean abundance of first and second generational peaks for *C. pulicaris* and *C. impunctatus* across Scotland in (A) 2008 and (B) 2009. Box plots show the median (central line), box denotes 25th and 75th percentiles, error bars represent 10th and 90th percentiles and dots are points outside the 10th and 90th percentiles.

significant positive association was confirmed in the linear mixed model (table 5). The same positive correlation between the size of the first and second peaks was found for *C. impunctatus* (2008 Pearson's $r=0.73$, $t_{14}=3.95$, $P=0.0014$; 2009 Pearson's $r=0.65$, $t_{10}=2.69$, $P=0.023$). Similarly, we included the size of the first peak in the best-fitting model identified during environmental variable selection. This confirmed the significant positive correlation between the size of the first and second peaks but, interestingly, did not improve model fit over the best-fitting model, including only environmental variables (table 5).

For *C. impunctatus*, three models had delta DIC values of less than two, including combinations of June and July temperatures and percent cover of heathland (table 5). Estimates for both of these variables were positive (table 5); however, none of these variables showed a significant relationship with the size of the second generational peak.

Model fit for the size of the first peak for *C. pulicaris* was poor ($R^2=0.34$, $F_{2,24}=7.7$, $P=0.0027$) but was very good for *C. impunctatus* ($R^2=0.97$, $F_{2,20}=383.3$, $P<0.001$). Model fit for the size of the second peak was good for both *C. pulicaris*

($R^2=0.74$, $F_{2,17}=28.2$, $P<0.001$) and for *C. impunctatus* ($R^2=0.98$, $F_{2,6}=292$, $P=0.001$).

Difference in size of first and second peaks

Three models for *C. pulicaris* had delta DIC values of less than two, including percent cover of pasture and agricultural land; however, neither of these variables showed significant relationships with the difference in size between the first and second peaks (table 6). It was not possible to separate the importance of the three variables selected to assess the difference in size of peaks for *C. impunctatus* because none of the variables showed a significant relationship with the difference in size of peaks, and all three variables were selected in the best models, all of which had delta DIC values of less than two (table 6).

Model fit for the difference in size of peaks for *C. pulicaris* was reasonable ($R^2=0.57$, $F_{2,17}=13.4$, $P<0.001$) but was very good for *C. impunctatus* ($R^2=0.99$, $F_{2,6}=1496$, $P<0.001$). Model fit for the size of the second peak was good for both *C. pulicaris* ($R^2=0.74$, $F_{2,17}=28.2$, $P<0.001$), and for *C. impunctatus* ($R^2=0.98$, $F_{2,6}=292$, $P=0.001$).

Mean parous rate

The mean parous rate for *C pulicaris* in 2008 was 0.33 (0.17 SD), which increased to 0.51 (0.22 SD) in 2009. *Culicoides impunctatus* has a mean parous rate 0.47 (0.22 SD) in 2008 and 0.42 (0.33 SD) in 2009.

Discussion

Insect abundance data presents challenges to statistical analysis, yet linking seasonal patterns of insect abundance to environmental drivers is critical for effective management of pests and disease. Researchers are often faced with supplying information to policy makers based upon short time-series of strongly fluctuating abundance data for species for which little ecological information is available. In this study, we have presented a methodology for linking environmental drivers with key phenological patterns of two Palearctic *Culicoides* species, using weekly abundance data collected from over just two years of observation. Importantly, the use of GAMs dealt with the large seasonal fluctuations in abundance exhibited by both of these species (coefficient of variation ∼7). Using a combination of smoothing functions and generalised linear mixed models, we identified critical timing of abundance peaks and identified correlations between abundance during these peaks and environmental drivers widely available from satellite and modelled temperature, precipitation and land-cover data. Model fit for most phenological traits was good, particularly for *C. impunctatus*, for which all best-fitting models had adjusted $R^2>0.6$ (four of the five models had $R^2>0.80$). However, model fit for *C. pulicaris* was less good, especially for the timing of the second peak and the size of the first peak. This may be because the resolution of our environmental variables was insufficiently fine to detect correlations or because important driving variables were not included in the modelling framework, such as host density on surrounding farmland. Traps were only run for one night per week, which may also have led to spurious variation in catches

Table 4. Model results for the size of the first peak in abundance for *C. pulicaris* and *C. impunctatus* in Scotland during 2008 and 2009. Posterior means for variable intercepts (*a*) for 2008 and 2009, and slope parameters for environmental variables (*b_i*) are shown with associated 95% credible intervals. The estimated number of parameters (*pD*), Deviance Information Criterion (DIC), and Gelmin-Rubin statistic for MCMC convergence (multivariate psrf) are shown for each model. Asterix (*) indicate significance for environmental variables. Δ_{DIC} refers to the difference in DIC values between each model and the model receiving the greatest support in the data. Dotted line separates model with approximately equal support ($\Delta_{DIC} < 2$) from those with considerably less support.

Species	Model	a 2008 a 2009	b_i	pD	DIC	Multivariate psrf	Δ_{DIC}
C. pulicaris	Degree days + spring ppt + pasture	9.99 (6.30,13.61) 8.99 (5.31,12.67)	−0.049 (−0.10,0.0035) −0.012 (−0.027,0.0020) 1.48 (−0.42,3.41)	8.8	208.5	1.03	0
	Degree days + spring ppt + pasture + spring temp	9.13 (4.69,13.47) 7.91 (3.06,12.69)	−0.066 (−0.14,0.0039) −0.013 (−0.027,0.0017) 1.48 (−0.46,3.43) 0.15 (−0.25,0.55)	9.8	209.9	1.02	1.4
	Degree days + spring ppt + pasture + spring temp + agriculture	9.29 (4.73,13.85) 8.08 (3.15,13.01)	−0.066 (−0.14,0.0035) −0.013 (−0.029,0.0021) 1.50 (−0.46,3.43) 0.15 (−0.26,0.57) −0.57 (−4.00,2.97)	10.8	211.8	1.02	3.3
	Degree days + spring ppt + pasture + spring temp + agriculture + Photoperiod	3.48 (−9.15,16.17) 2.18 (−10.68,15.13)	−0.064 (−0.14,0.0090) −0.010 (−0.027,0.0059) 1.92 (−0.23,3.99) 0.21 (−0.21,0.64) −0.85 (−4.15,2.54) 0.15 (−0.15,0.45)	11.6	212.6	1.02	4.1
	Degree days + spring ppt + pasture + spring temp + agriculture + Photoperiod + winter temp	4.63 (−9.36,18.34) 3.22 (−10.81,16.97)	−0.062 (−0.13,0.012) −0.012 (−0.029,0.0061) 1.91 (−0.28,4.02) 0.20 (−0.23,0.63) −0.73 (−4.25,2.81) 0.12 (−0.19,0.45) −0.11 (−0.68,0.46)	12.5	214.6	1.02	6.1
C. impunctatus	Photoperiod + degree days + spring precipitation	−12.28 (−29.26,4.06) −12.13 (−29.22,4.24)	0.46 (0.02,0.91)* 0.060 (−0.018,0.13) 0.020 (−0.0021,0.041)	21.7	166.5	1.02	0
	Photoperiod + degree days + spring precipitation + heath	−12.39 (−31.26,6.36) −12.24 (−31.19,6.54)	0.46 (−0.052,0.98) 0.061 (−0.020,0.14) 0.020 (−0.0041,0.042) −0.067 (−6.78,6.54)	22.1	167.1	1.02	0.6
	Photoperiod + degree days + spring precipitation + heath + spring temp	−9.22 (−28.85,10.11) −8.79 (−28.65,10.78)	0.42 (−0.092,0.94) 0.077 (−0.0078,0.16) 0.019 (−0.0038,0.042) 0.035 (−6.56,6.29) −0.23 (−0.68,0.19)	22.4	169.7	1.02	3.2
	Photoperiod + degree days + spring precipitation + heath + spring temp + winter temp	−6.46 (−26.37,12.93) −6.35 (−26.41,13.12)	0.38 (−0.12,0.89) 0.075 (−0.011,0.16) 0.018 (−0.0053,0.041) −1.06 (−7.62,5.76) −0.28 (−0.75,0.16) −0.36 (−1.09,0.33)	22.8	172.5	1.02	6.0

caused by local microclimatic effects, for which we were unable to account in our models.

The timing of generational peaks identified using our modelling approach was reasonably consistent with previous empirical data on other UK populations of *C. pulicaris* and *C. impunctatus*. Emergence in spring has been observed to begin in late April/early May in *C. pulicaris* (Service, 1974; Holmes & Boorman, 1987) and early May in *C. impunctatus* (Blackwell *et al.*, 1992), which is consistent with our finding that the first generation of *C. pulicaris* peaks earlier than that of *C. impunctatus*.

For both species, our finding that the first generational peak was on average larger than the second was also consistent with previous studies in southern England, where the greatest numbers of adult *C. pulicaris* were caught in June (Service, 1974) or July (Service, 1968) and those of *C. impunctatus* in June (Hill, 1947; Service, 1969; Boorman & Goddard, 1970) or July (Parker, 1949).

Table 5. Model results for the size of the second peak in abundance for *C. pulicaris* and *C. impunctatus* in Scotland during 2008 and 2009. Posterior means for variable intercepts (*a*) for 2008 and 2009, and slope parameters for environmental variables (*b$_i$*) are shown with associated 95% credible intervals. The estimated number of parameters (*pD*), Deviance Information Criterion (DIC), and Gelmin-Rubin statistic for MCMC convergence (multivariate potential scale reduction factor) are shown for each model. Asterix (*) indicate significance for environmental variables. Δ$_{DIC}$ refers to the difference in DIC values between each model and the model receiving the greatest support in the data. Dotted line separates model with approximately equal support (Δ$_{DIC}$ < 2) from those with considerably less support.

Species	Model	*a* 2008 *a* 2009	*b$_i$*	*pD*	DIC	Multivariate psrf	Δ$_{DIC}$
C. pulicaris	Pasture + agriculture + size of first generational peak	−2.29 (−6.80,2.10) −1.07 (−4.79,2.45)	2.30 (0.11,4.46)* 2.38 (−1.44,6.11) 0.47 (0.13,0.83)*	11.9	166.3	1.02	0
	Pasture + agriculture pasture	3.33 (1.86,4.82) 3.34 (1.88,4.86) 3.68 (2.22,5.08) 3.72 (2.30,5.15)	2.97 (0.73,5.15)* 2.81 (−1.07,6.80) 2.70 (0.51,4.89)*	10.7 10.4	171.0 171.3	1.02 1.03	4.7 5.0
	Pasture + agriculture + June and July ppt	4.21 (0.53,7.98) 4.14 (0.70,7.64)	3.32 (0.77,5.94)* 2.62 (−1.53,6.65) −0.0055 (−0.026,0.014)	11.8	172.4	1.02	6.1
	Pasture + agriculture + June and July ppt + June and July temp	−0.24 (−8.34,7.85) −0.55 (−9.02,7.77)	2.99 (0.44,5.59)* 2.90 (−1.33,7.07) −0.0038 (−0.025,0.017) 0.26 (−0.17,0.69)	12.7	172.6	1.02	6.3
C. impunctatus	June and July temp	4.13 (−2.21,10.90) 3.74 (−3.00,10.92)	0.17 (−0.24,0.56)	15.3	85.1	1.02	0
	June and July temp + heath	4.29 (−2.15,11.20) 3.90 (−2.74,11.32)	0.15 (−0.28,0.54) 2.94 (−1.67,7.44)	15.0	85.9	1.02	0.8
	Heath	6.66 (5.79,7.54) 6.42 (5.44,7.34)	3.14 (−1.47,7.87)	14.3	86.5	1.02	1.4
	June and July ppt	7.12 (3.42,10.82) 6.90 (3.25,10.54)	−0.0012 (−0.020,0.017)	15.3	87.5	1.02	2.4
	June and July temp + heath + June and July ppt	4.07 (−3.95,12.21) 3.69 (−4.65,12.14)	0.14 (−0.30,0.55) 3.06 (−1.74,7.80)	15.4	88.2	1.02	3.1
	Heath + June and July ppt	6.44 (2.62,10.22) 6.20 (2.43,9.89)	3.14 (−1.56,7.88) 0.0011 (−0.017,0.020)	14.7	88.6	1.02	3.5
	June and July temp + size of the first generational peak	−3.79 (−12.16,3.63) −4.10 (−12.73,3.59)	0.16 (−0.21,0.56) 0.55 (0.26,0.89)	11.02	89.2	1.02	4.1

The positive association between timing of the first peak and the day of the year at which eight hours of daylight was passed in both species corroborates previous studies that have documented earlier emergence at lower latitudes within the UK (southern England: April/May: Service, 1968, 1969, 1974; Boorman & Goddard, 1970; Holmes & Boorman, 1987; Scotland: May/June: Blackwell *et al.*, 1992). At present, we can only speculate on the underlying causes for the links this study identified between environmental variables and patterns of timing. Our finding that *C. impunctatus* sites experiencing warmer winters tended to have later first peaks in abundance may perhaps be related to a less synchronous overwintering period, and it is possible that drier conditions in the summer caused *C. pulicaris* sites experiencing warmer summers to have later second generational peaks. However, further study is required to more mechanistically link these associations with underlying ecological processes. The negative correlation identified in 2008 between the timing of the first and second peaks for *C. pulicaris* suggests that in this year, sites with earlier first peaks tended to have later second peaks, and vice versa. This may be reflective of sites with environmental conditions favourable for earlier emergence also tending to have conditions that allow for a longer season of activity or at least a wider second peak. However, it is difficult to interpret this finding with only two years of data; and,

indeed, in the second year of study, this positive relationship was not detected. No significant correlations in the timing of generational peaks were identified for *C. impunctatus*, suggesting that for this species, the emergence of the second generation may be less strongly linked to that of the first generation than in *C. pulicaris*; but, again, this is very difficult to interpret with only two years of data.

Environmental factors governing the size of abundance peaks not only differed between species but also between the generations of a species, suggesting that the influence, particularly of climatic factors, can alter throughout the year. Given the life cycle of these palearctic species, abundance during the first generation is likely to be related to the suitability of conditions for survival and development of larvae overwinter from diapause, and the survival and rate of completion of the adult reproduction cycle. Similarly, abundance of the second generation is likely to be linked with survival and development of the larvae from eggs laid by the first generation (and, therefore, also to the size of the first generation) and to the survival and rate of completion of the adult reproductive cycle. No significant effects on the size of the first generational peak were detected for *C. pulicaris*, and only photoperiod was found to significantly impact the size of the first generational peak for *C. impuncatus*, with sites at higher latitudes tending to have larger abundances, possibly

Table 6. Model results for the difference in size between the first and second peaks in abundance (log [peak 1/peak 2]) for *C. pulicaris* and *C. impunctatus* in Scotland during 2008 and 2009. Posterior means for variable intercepts (*a*) for 2008 and 2009, and slope parameters for environmental variables (b_i) are shown with associated 95% credible intervals. The estimated number of parameters (pD), Deviance Information Criterion (DIC), and Gelmin-Rubin statistic for MCMC convergence (multivariate psrf) are shown for each model. Δ_{DIC} refers to the difference in DIC values between each model and the model receiving the greatest support in the data. No environmental variables were significant. Dotted line separates model with approximately equal support ($\Delta_{DIC} < 2$) from those with considerably less support.

Species	Model	a2008 a 2009	b_i	pD	DIC	Multivariate psrf	Δ_{DIC}
C. pulicaris	Pasture	1.78 (0.28,3.26) 0.41 (0.28,3.26)	−1.30 (−3.59,0.95)	8.1	177.8	1.03	0
	Agriculture	1.18 (0.22,20.9) −0.17 (−1.04,0.67)	1.38 (−5.38,2.64)	8.2	178.7	1.02	0.9
	Pasture + agriculture	2.00 (0.37,3.59) 0.63 (−0.93,2.19)	−1.46 (−3.81,0.85) −1.78 (−5.90,2.33)	8.9	178.7	1.03	0.9
	Pasture + agriculture + June and July temp	4.69 (−2.68,12.07) 3.53 (−4.33,11.37)	−1.32 (−3.61,1.03) −1.96 (−6.03,2.07) −0.17 (−0.62,0.28)	9.7	180.4	1.02	2.6
	Pasture + agriculture + June and July temp + June and July ppt	5.46 (−3.42,14.34) 4.24 (−4.96,13.41)	−1.03 (−3.79,1.66) −2.21 (−6.48,2.07) −0.17 (−0.66,0.29) −0.0044 (−0.027,0.018)	10.7	182.3	1.02	4.5
C. impunctatus	Heath	0.17 (−0.44,0.97) 0.10 (−0.66,0.92)	−0.27 (−4.33,3.76)	9.1	103.1	1.02	0
	June and July temp	3.87 (−3.95,11.96) 3.90 (−4.14,12.23)	−0.22 (−0.71,0.24)	8.3	103.3	1.02	0.2
	June and July ppt	1.37 (−2.53,5.11) 1.24 (−2.54,4.85)	−0.0063 (−0.026,0.014)	8.1	103.6	1.02	0.5
	Heath + June and July temp	3.77 (−4.46,12.11) 3.80 (−4.68,12.44)	−0.055 (−4.14,4.02) −0.22 (−0.72,0.28)	9.4	104.7	1.02	1.6
	Heath + June and July ppt	1.45 (−2.65,5.53) 1.32 (−2.87,5.27)	−0.46 (−4.53,3.80) −0.0066 (−0.027,0.015)	9.3	105.0	1.02	1.9
	June and July temp + June and July ppt	5.65 (−3.80,15.10) 5.63 (−3.96,15.34)	−0.24 (−0.74,0.25) −0.0078 (−0.028,0.013)	8.4	105.0	1.02	1.9
	Heath + June and July temp + June and July ppt	5.59 (−4.08,15.29) 5.57 (−4.28,15.43)	−0.34 (−4.31,3.73) −0.24 (−0.76,0.27) −0.0078 (−0.028,0.014)	9.2	106.8	1.02	3.7

reflecting the more northerly distribution of this species. The positive impact of the percent cover of pasture on abundance during the second generational peak for *C. pulicaris* is consistent with previous studies for this livestock-associated species (Takken *et al.*, 2008). The positive correlation between the size of the first and second peaks for *C. pulicaris* and *C. impunctatus* detected in both years is likely reflective of sites with favourable environmental conditions tending to produce larger abundances during both generational peaks. Including the size of the first peak in the best-fitting linear mixed model for *C. pulicaris* improved model fit substantially, but did not change the magnitude or significance of the effects of environmental variables, lending support to the robustness of the positive association between percent cover of pasture and the size of the second peak for this species. However, for *C. impunctatus*, including the size of the first peak in the best-fitting linear mixed model did not improve model fit over models including only environmental variables. Whether this result is indicative of a stronger effect of exogenous forces on abundance compared to intrinsic population dynamics in this species needs to be confirmed by further research. We detected no significant correlations for the difference in size between the first and second peaks in abundance for either species, suggesting that the environmental variables used in this analysis did not adequately capture the processes affecting

mortality between peak abundances. It is also likely that the relatively low occurrence of heathland around the trapping sites meant that we sampled from relatively sub-optimal habitat for *C. impunctatus* populations. This may have been responsible for the lack of correlation between landcover variables and population size in this species.

The likely mechanisms underlying the variation between species and generations in observed climatic responses are unclear, due to the lack of life-history data for Palearctic species to assist in interpretation. Temperature has been found, largely from laboratory studies on temperate North American colony species, to have opposing effects on different stages of *Culicoides* life cycles. There are negative effects on fecundity (Linley & Hinds, 1976; Kramer *et al.*, 1985), and both adult (Hunt *et al.*, 1989; Wellby *et al.*, 1996; Wittmann, 2000) and larval survival rates, but positive effects on adult activity rates (e.g. mating, feeding) (Blackwell, 1997; Kettle *et al.*, 1998), rates of larval development and completion of the gonado-trophic cycle (in terms of both oocyte development and subsequent egg deposition (Linley, 1966; Akey *et al.*, 1978; Edwards, 1982; Kitoaka, 1982; Mullens & Rutz, 1983; Allingham, 1991; Mullens & Holbrook, 1991; Bishop *et al.*, 1996; Wittman, 2000)).

We were unable to detect any influence of precipitation on the phenological traits for either species, perhaps because the

resolution of the precipitation data (25 km^2) was too coarse. The influence of precipitation on life history characteristics of *Culicoides* is not well documented; however, precipitation has been linked to the successful development of higher numbers of larvae (Mellor *et al.*, 2000) due to the greater availability and permanence of semi-aquatic breeding sites.

The mean parous rate is an estimate of the survival rate per gonotrophic cycle, and the values deduced from this study for *C. impunctatus* in 2008 (0.47) and 2009 (0.42) are similar to that from a study in Argyllshire in southeast Scotland (0.57: Blackwell *et al.*, 1992). The parity rate averaged across all sites for *C. pulicaris* was considerably lower in 2008 (0.33) than in 2009 (0.51), suggesting a degree of sensitivity to environmental conditions (i.e. the warmer spring and drier summer in 2009). The values for *C. pulicaris* from this study during 2009 are relatively high compared to the range of parity rate estimates for other temperate *Culicoides* species: 0.61 and 0.24 for *C. variipennis* in New York State (Mullens & Rutz, 1984) and California (Work *et al.*, 1991), and 0.38 for *C. obsoletus* in the UK (Holmes & Birley, 1987). We detected a third generational peak for several sites: five sites in 2008 and four sites in 2009 for *C. pulicaris*, and three sites with a third generation in 2008, and one site in 2009 for *C. impunctatus*. The third peaks were generally small in 2008 (range 10–334) but were considerably larger in 2009 at some sites (range 7–8671). These findings are suggestive of trivoltinism at some trapping locations for both *C. pulicaris* and *C. impunctatus*. Although we could find no records in the literature of trivoltinism for these species, a study in southern England suggested that *C. obsoletus*, a related temperate species, appears to be trivoltine (Birley & Boorman, 1982). This variation warrants further studies with daily trapping data to confirm voltinism in these species, particularly because of the potential influence a third abundance peak could have on disease transmission late in the vector season.

In summary, the techniques implemented in this study provide a framework for analysing short time-series of insect count data, dealing with widely fluctuating data containing a large proportion of zeros. Applying these techniques to a longer time-series of data would allow for analysis of environmental drivers of additional phenological traits such as overwintering mortality in temperate species, an area currently lacking in empirical study. Extending this approach to the analysis of insect count data across larger spatial and temporal scales is easily achieved, following well-established hierarchical GLMM methods for grouping data across space, time and species (e.g. Gelman & Hill, 2007). Though our current explanations of the mechanisms underlying the differential environmental responses of the species studied are somewhat speculative, this underscores the need for a greater understanding of the effects of temperature and precipitation on the life history traits of palearctic *Culicoides* spp. in Europe. Empirical studies that contrast the environmental responses of different *Culicoides* spp. within the same model framework are invaluable for enhancing this understanding, allowing season- and species-specific impacts to be detected and related to life history characteristics.

Acknowledgements

References

Akey, D.H., Potter, H.W. & Jones, R.H. (1978) Effects of rearing temperature and larval density on longevity, size and fecundity in the biting gnat *Culicoides* viriipennis. *Annals of the Entomological Society of America* **71**, 411–418.

Allingham, P. (1991) Effect of temperature on late immature stage of *Culicoides* brevitarsis (Diptera: Ceratopogonidae). *Journal of Medical Entomology* **28**, 878–881.

Baylis, M., Mellor, P.S., Wittmann, E.J. & Rogers, D.J. (2001) Prediction of areas around the Mediterranean at risk of bluetongue by modelling the distribution of its vector using satellite imaging. *Veterinary Record* **149**, 639–643.

Birley, M.H. & Boorman, J.P.T. (1982) Estimating the survival and biting rates of haematophagus insects, with particular reference to the *Culicoides obsoletus* group (Diptera: Ceratopogonidae) in southern England. *Journal of Animal Ecology* **51**, 135–148.

Bishop, A.L., McKenzie, H.J., Barchia, I.M. & Harris, A.M. (1996) Effect of temperature regimes on the development, survival and emergence of *Culicoides* brevitarsis Kieffer (Diptera: Ceratopogonidae). *Australian Journal of Entomology* **35**, 361–168.

Blackwell, A. (1997) Diel flight periodicity of the biting midge *Culicoides impunctatus* and the effects of meteorological conditions. *Medical and Veterinary Entomology* **11**, 361–367.

Blackwell, A., Mordue, A.J., Young, M.R. & Mordue, W. (1992) Bivoltinism, survival rates and reproductive characteristics of the Scottish biting midge, *Culicoides impunctatus*, (Diptera: Ceratopogonidae) in Scotland. *Bulletin of Entomological Research* **82**, 299–306.

Boorman, J. (1986) British *Culicoides* Diptera Ceratopogonidae notes on distribution and biology. *Entomologist's Gazette* **37**, 253–266.

Boorman, J. & Goddard, P. (1970) Observations on the biology of *Culicoides impunctatus* Goetgh. (Dipt., Ceratopogonidae) in southern England. *Bulletin of Entomological Research* **60**, 189–198.

Braverman, Y. & Mumcuoglu, K.Y. (2009) Newly emerged nulliparous *Culicoides imicola* Kieffer (Diptera: Ceratopogonidae) with pigmented abdomen. *Medical and Veterinary Entomology* **160**, 356–358.

Campbell, J.A. & Pelham-Clinton, E.C. (1960) A taxonomic review of the British species of *Culicoides* Latreille (Diptera: Ceratopogonidae). *Proceedings of the Royal Society of Edinburgh B, LXVII* **21**, 181–301.

Caracappa, S., Torina, A., Guerico, A., Vitale, F., Calabro, A., Ferrantelli, G., Vitale, M. & Mellor, P.S. (2003) Identification of a novel bluetongue virus vector species of *Culicoides* in Sicily. *Veterinary Record* **153**, 71–74.

Dyce, A.L. (1969) The recognition of nulliparous and parous *Culicoides* (Diptera: Ceratopoginidae) without dissection. *Australia Journal of Entomology* **8**, 11–15.

Eber, S. (2004) Bottom-up density regulation in the holly leaf-miner *Phytomyza ilicis*. *Journal of Animal Ecology* **73**, 948–958.

Edwards, P.B. (1982) Laboratory observations on the biology and life cycle of the Australian biting midge *Culicoides subimmaculatus* (Diptera: Ceratopogonidae). *Journal of Medical Entomology* **19**, 545–552.

EEA (European Environment Agency) (2000) Corine Landcover 2000: mapping a decade of change. Copenhagen, Denmark, European Environment Agency.

Gelman, A. & Hill, J. (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York, USA, Cambridge University Press.

Gelman, A. & Rubin, D.B. (1992) Inference from iterative simulation using multiple sequences. *Statistical Science* **7**, 457–511.

Gern, L., Cadenas, F.M. & Burri, C. (2008) Influence of some climatic factors on *Ixodes ricinus* ticks studied along altitudinal gradients in two geographic regions in Switzerland. *International Journal of Medical Microbiology* **298**, 55–59.

Haylock, M.R., Hofstra, N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D. & New, M. (2008) A European daily high-resolution gridded dataset of surface temperature and precipitation. *Journal of Geophysical Research (Atmospheres)* **113**, 1–12.

Heffernan, J.M., Smith, R.J. & Wahl, L.M. (2005) Perspectives on the basic reproductive ratio. *Journal of the Royal Society Interface* **2**, 281–293.

Hendry, G. & Goodwin, G. (1988) Biting midges in Scottish forestry: a costly irritant or a trivial nuisance? *Scottish Forestry* **42**, 113–119.

Hill, M.A. (1947) The lifecycle and habits of *Culicoides impunctatus* Goetghebuer and *Culicoides obsoletus* Meigen, together with some observations on the lifecycle of *Culicoides pallidicornis* Kieffer, *Culicoides cubitalis* Edwards and *Culicoides chiopterus* Meigen. *Annals of Tropical Medicine and Parasitology* **41**, 55–115.

Hodgson, J.A., Thomas, C.D., Oliver, T.H., Anderson, B.J., Brereton, T.M. & Crones, E.E. (2010) Predicting insect phenology across space and time. *Global Change Biology* **17**, 1289–1300.

Hoeting, J.A. (2009) The importance of accounting for spatial and temporal correlation in analyses of ecological data. *Ecological Applications* **19**, 574–577.

Holmes, P.R. & Birley, M.H. (1987) An improved method for survival rate analysis from time series of haematophagus Dipteran populations. *Journal of Animal Ecology* **56**, 427–440.

Holmes, P.R. & Boorman, P.T. (1987) Light suction trap catches of *Culicoides* midges in southern England. *Medical and Veterinary Entomology* **1**, 349–359.

Hunt, G.J., Tabachnick, W.J. & McKinnon, C.N. (1989) Environmental factors affecting mortality of adult *Culicoides variipennis* (Diptera: Ceratopogonidae) in the laboratory. *Journal of the American Mosquito Control Association* **5**, 387–391.

Institute for Animal Health (2011) *Culicoides*. NET: Incorporating the UK Culicoides reference laboratory. Available online at http://www.culicoides.net, (accessed 17 November 2011).

Kettle, D.S. (1950) The seasonal distribution of *Culicoides impunctatus* Goetghebuer (Diptera: Heleidae (Ceratopogonidae)) with a discussion on the possibility that it may be composed of two or more biological races. *Transactions of the Royal Society of London* **101**, 125–145.

Kettle, D.S., Edwards, P.B. & Barnes, A. (1998) Factors affecting numbers of *Culicoides* in truck traps in Coastal Queensland. *Medical and Veterinary Entomology* **12**, 367–377.

Kitoaka, S. (1982) Effects of rearing temperature on length of larval period and size in adults *Culicoides arakawae* and *Culicoides maculates*. *National Institute for Animal Health Quarterly (Japan)* **22**, 159–162.

Kramer, W.L., Greiner, E.C. & Gibbs, E.P.J. (1985) Seasonal variations in population size, fecundity, and parity rates of *Culicoides insignis* (Diptera: Ceratopogonidae) in Florida, USA. *Journal of Medical Entomology* **22**, 163–169.

Linley, J.R. (1966) The ovarian cycle in *Culicoides barbosia* Wirth and Blanton and *C. furens* (Poey) (Diptera: Ceratopogonidae). *Bulletin of Entomological Research* **57**, 1–17.

Linley, J.R. & Hinds, M.J. (1976) Seasonal changes in size, female fecundity and male potency in *Culicoides melleus* (Diptera: Ceratopogonidae). *Journal of Medical Entomology* **13**, 151–156.

Lord, C.C. & Baylis, M. (1999) Estimation of survival rates in haematophagus insects. *Medical and Veterinary Entomology* **13**, 225–233.

Lunn, D., Spiegelhalter, D., Thomas, A. & Best, N. (2009) The BUGS project: evolution, critique and future directions. *Statistics in Medicine* **28**, 3049–3067.

MacDonald, G. (1952) The objective of residual insecticide campaigns. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **46**, 227–235.

MacDonald, G. (1957) *Epidemiology and Control of Malaria*. London, UK, Oxford University Press.

McCarthy, M.A. (2007) *Bayesian Methods for Ecology*. Cambridge, UK, Cambridge University Press.

Mellor, P.S., Boorman, J. & Baylis, M. (2000) *Culicoides* biting midges: their role as arbovirus vectors. *Annual Review of Entomology* **45**, 307–340.

Mullens, B.A. & Holbrook, F.R. (1991) Temperature effects on the gonotrophic cycle of *Culicoides variipennis* (Diptera: Ceratopogonidae)). *Journal of the American Mosquito Control Association* **7**, 588–591.

Mullens, B.A. & Rutz, D.A. (1983) Development of immature *Culicoides variipennis* (Diptera: Ceratopogonidae) at constant laboratory temperature. *Annals of the Entomological Society of America* **76**, 71–76.

Mullens, B.A. & Rutz, D.A. (1984) Age structure and survivourship and *Culicoides variipennis* (Diptera: Cerapogonidae) in central New York State, USA. *Journal of Medical Entomology* **21**, 194–203.

Parham, P.E. & Michael, E. (2010) Modelling climate change and malaria transmission. *Modelling Parasite Transmission and Control* **673**, 184–199.

Parker, A.H. (1949) Observations on the seasonal and daily incidence of certain biting midges (*Culicoides* Latr.) in Scotland. *Transactions of the Royal Entomological Society London* **100**, 170–190.

Purse, B.V., Tatem, A.J., Caracappa, S., Rogers, D.J., Mellor, P.S., Baylis, M. & Torina, A. (2004) Modelling the distributions of *Culicoides* bluetongue virus vectors in Sicily in relation to satellite-derived climate variables. *Medical and Veterinary Entomology* **18**, 90–101.

R Development Core Team (2009) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Rawlings, P. (1996) A key, based on wing patterns of biting midges (genus *Culicoides* Latreille – Diptera: Ceratologonidae) in the Iberian penninsula, for use in epidemiological studies. *Graellsia* **52**(11), 57–71.

Rodeghiero, M. & Battisti, A. (2000) Inter-tree distribution of the spruce web-spinning sawfly, *Cephalcia abietis*, at endemic density. *Agricultural and Forest Entomology* **2**, 291–296.

Saldaña, S., Lima, M. & Estay, S. (2007) Northern Atlantic Oscillation effects on the temporal and spatial dynamics of green spruce aphid populations in the UK. *Journal of Animal Ecology* **76**, 782–789.

Service, M.W. (1968) Light-trap catches of *Culicoides* spp. (Dipt., Cerapogonoidae) from southern England. *Bulletin of Entomological Research* **59**, 317–322.

Service, M.W. (1969) Studies on the biting habits of *Culicoides impunctatus* Goetghebuer, *C. obsoletus* (Meigen) and *C. punctatus* (Meigen) (Dipetera: Ceratopogonidae) in southern England. *Proceedings of the Royal Entomological Society London, Series A* **44**, 1–10.

Service, M.W. (1974) Further results of catches of *Culicoides* (Dipetra: Ceratopogonidae) and mosquitos from suction traps. *Journal of Medical Entomology* **11**, 471–479.

Sileshi, G. (2006) Selecting the right statistical model for analysis of insect count data by using information theoretic measures. *Bulletin of Entomological Research* **96**, 479–488.

Speigelhalter, D.J., Best, N.G., Carlin, B.P. & Van Der Linde, A. (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)* **64**, 583–639.

Takken, W., Verhulst, N., Scholte, E.J., Jacobs, F., Jongema, Y. & van Lammeren, R. (2008) The phenology and population dynamics of *Culicoides* spp. in different ecosystems in the Netherlands. *Preventative Veterinary Medicine* **87**, 41–54.

Tatem, A.J., Baylis, M., Mellor, P.S., Purse, B.V., Capela, R., Pena, I. & Rogers, D.J. (2003) Prediction of bluetongue vector distribution in Europe and north Africa using satellite imagery. *Veterinary Microbiology* **97**, 13–29.

Van Ark, H. & Meiswinkel, R. (1992) Subsampling of large light-trap catches of *Culicoides* (Diptera: Ceratopogonidae). *Onderstepoort Journal of Veterinary Research* **59**, 183–189.

Vinatier, F., Tixier, P., Duyck, P. & Lescourret, F. (2011) Factors and mechanisms explaining spatial heterogeneity: a review of methods for insect populations. *Methods in Ecology and Evolution* **2**, 11–22.

Wan, Z. & Li, Z.L. (1997) A physics-based algorithm for retrieving land-surface emissivity and temperature from EOS/MODIS data. *IEEE Transactions on Geoscience and Remote Sensing* **35**, 980–996.

Wellby, M.P., Baylis, M., Rawlings, P. & Mellor, P.S. (1996) Effect of temperature on virogenesis of African horse sickness virus in *Culicoides variipensis sonorensis* (Diptera: Ceratopogonidae) and its significance in relation to the epidemiology of the disease. *Bulletin of Entomological Research* **86**, 715–720.

Wittmann, E.J. (2000) Temperature and transmission of arboviruses by *Culicoides* biting midges. PhD thesis, University of Bristol, Bristol, UK.

Wood, S.N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society, Series B* **73**, 3–36.

Work, T.M., Mullens, B.A. & Jessup, D.A. (1991) Estimation of survival and gonotrophic cycle length of *Culicoides variipennis* (Diptera: Cerapogonidae) in California. *Journal of the American Mosquito Control Association* **7**, 242–249.

Yang, G., Brook, B.W., Whelan, P.I., Cleland, S. & Bradshaw, C.J. A. (2008) Endogenous and exogenous factors controlling temporal abundance patterns of tropical mosquitoes. *Ecological Applications* **18**, 2028–2040.

Yang, G., Brook, B.W. & Bradshaw, C.J.A. (2009) Predicting the timing and magnitude of tropical mosquito population peaks for maximising control efficiency. *PLOS Neglected Tropical Diseases* **3**, 1–9.