**Supplemental material**

Ruling out deer density as a driver of hemorrhagic disease patterns in the southeast United States

We collated categorical data of white-tailed deer density estimates collected by state departments of natural resources (DNR). The categories recorded are 0: no deer, 1: less than 15 deer/mile2, 2: between 15 & 30 deer/mile2, 3: between 30 & 45 deer/mile2, and 4: more than 45 deer/mile2. Some DNR agencies record density at the county level, while others use a natural landscape. For consistency, we imported graphic versions of data into ArcGIS, and rectified it onto a standard polygon layer of the US counties in the WGS84 geodetic system and geographic coordinates. We then converted and exported the rectified image into a raster layer of the entire lower 48 states. Finally, we then imported this raster layer in the R programming environment using the ‘raster’ package, and extracted the modal density category from the pixels of each county. The result is a tabular dataset of deer density estimates for each county. We linked this dataset with the existing disease dataset, scoring each county according to the number of years in 2008-2012 that it recorded hemorrhagic disease (0-5). We considered all the counties of the six study states (AR, LA, MS, AL, GA, FL) and found no relationship between deer density category and number of years reporting hemorrhagic disease (correlation tests: Pearson’s **=0.03, *p*=0.45; Spearman’s **=0.05, *p*=0.26).

Estimating *Culicoides* community activity periods across sites

For each species, the probability of detection in each week of the year is calculated (from the temporally aggregated light trap data) as the proportion of light traps set in a given week of the year that were found to contain the target species. As these probabilities are approximately normally distributed, we can determine a range around the week associated with the peak probability of detection (the mean, or peak, detection week calculated by a weighted sum). Our choice of all weeks within 80% of the peak detection week is ascertained by noting that the probability density function for the standard normal *N*(0,1) is given by , maximized at  to the value . If we want to define a range around the mean such that it includes only values 80% of the maximum then we usewhich is solved by , i.e., 66% of the standard deviation (s.d.) around the mean (means and 0.66 x s.d. shown in Fig. A1). The community activity period, **, is then the proportion of weeks of the year represented in the combined temporal ranges of each member species of a community. This allows community phenology to reflect both peak abundance weeks for member species but also the variation around that peak. The 80% threshold is also conservative in the sense that it does not reflect early emergence and late overwintering observed only at relatively southern latitudes, since these time periods are in the tails of the distributions and typically correspond to a detection probability <<80% of a species’ maximum.



Fig. A1: Weeks of the year in which *Culicoides* species are likely to be active, as determined by temporally stratified light trap data (2008-2012). Circles represent mean detection week and vertical lines show 66% of the standard deviation, which corresponds to a species detection probability of 80% of its maximum

Modeling vectors in active and inactive states

The mathematical model introduced in the main text allows the vector population to switch in and out of the active state, spending a fraction of the year ** in the active state (Fig. A2). The rate of switching is determined by the size of the *f* parameters.

****

Fig. A2: Examples of vectors switching between active and inactive phases for **=0.5 (top panel) and **=0.8 (bottom panel). The switching parameters ( and ) alternately take the values 0.0 and 100.0.

Derivation of *R*0

For the model with continuously active vectors (**=1.0), we identify two infectious classes, *Iv* and *Ih*. The disease-free equilibrium is. The gains to and losses from the infectious classes are summarized as

|  |  |
| --- | --- |
| Gains to *Iv* |  |
| Gains to *Ih* |  |
| Losses from *Iv* |  |
| Losses from *Ih* |  |

The *F* (gains) and *V* (losses) matrices are given by

 and

which simplify to

 and 

The matrix *G*=*FV*-1 is given by

*G*

The largest eigenvalue of *G*, which is the basic reproductive number, is then



Relationship between vector species richness and abundance

The relationship between vector species richness at a site and average abundance (total abundance/number of light traps used at a site) increases between species richness measures of 1-5 after which abundance begins to asymptote to a maximum value at higher measures of species richness (Fig. A3). This suggests that interspecific competition may limit the overall abundance of *Culicoides* communities.



Fig. A3: Average community abundance at each site as a function of species richness. A trend line (loess, span=0.8) is fitted and the shaded region around the line represents 95% confidence intervals calculated using the *t*-based approximation.

Some *Culicoides* species have earlier emergence and later overwintering at southern latitudes

To determine potentially confounding effects of latitude on phenology patterns, the *Culicoides* data were divided into two halves (at the approximate latitudinal midpoint: 30 degrees latitude, Fig. 1), which maintained balance in the light trap data (*n*north=2618, *n*south=2660). This latitudinal stratification reveals that for 11 of the 52 species, the distribution of their observation week (1-52; aggregated over the 5 sampling years) is different in the north versus the south (Fig. A4, Kolmogorov-Smirnov (KS) tests: p<0.05 – red=south, blue=north). For the other species the KS test was either not significant (Fig. A4, pink=south and cyan=north symbols, with black outline), or not performed due to the species occurring exclusively in one of the two latitudinal groups (Fig. A4, pink and blue symbols without black outline). The general picture that emerges among the species whose phenology does differ latitudinally is earlier emergence and later overwintering in the south. These observations support the rationale to define activity periods as the interval when detection probability is >80% of its maximum as it conservatively estimates species-level activity region-wide. Importantly, extended activity is not in disease hotspot regions, meaning that disease patterns are not driven by *Culicoides* activity at the species level (rather at the community level, Fig. 1).



Fig. A4: Weeks of the year (2008-2012 aggregated) in which each species was detected in the southern region (<30 degrees latitude; red/pink symbols) and northern region (>30 degrees latitude; blue/cyan symbols). Kolmogorov-Smirnov tests were performed on species whose symbols have a black outline, with significantly different distributions colored red/blue (versus pink/cyan). Species whose symbols do not have a black outline occur exclusively in one of the two regions, and consequently the test was not performed.