

# A new approach to interspecific synchrony in population ecology using tail association

Shyamolina Ghosh<sup>1</sup>  | Lawrence W. Sheppard<sup>1</sup>  | Philip C. Reid<sup>2,3</sup> | Daniel Reuman<sup>1</sup> 

<sup>1</sup>Department of Ecology and Evolutionary Biology and Kansas Biological Survey, University of Kansas, Lawrence, KS, USA

<sup>2</sup>Continuous Plankton Recorder Survey, The Marine Biological Association, The Laboratory, Plymouth, UK

<sup>3</sup>School of Biological & Marine Sciences, University of Plymouth, Plymouth, UK

## Correspondence

Daniel Reuman, University of Kansas, 2101 Constant Ave, Lawrence, KS 66047, USA.  
Email: reuman@ku.edu

## Funding information

James S McDonnell Foundation.; Biotechnology and Biological Sciences Research Council, Grant/Award Number: BBS/E/C/000J0200; US National Science Foundation, Grant/Award Number: 1714195 and 1442595; James S McDonnell Foundation

## Abstract

Standard methods for studying the association between two ecologically important variables provide only a small slice of the information content of the association, but statistical approaches are available that provide comprehensive information. In particular, available approaches can reveal *tail associations*, that is, accentuated or reduced associations between the more extreme values of variables. We here study the nature and causes of tail associations between phenological or population-density variables of co-located species, and their ecological importance. We employ a simple method of measuring tail associations which we call the *partial Spearman correlation*. Using multidecadal, multi-species spatiotemporal datasets on aphid first flights and marine phytoplankton population densities, we assess the potential for tail association to illuminate two major topics of study in community ecology: the stability or instability of aggregate community measures such as total community biomass and its relationship with the synchronous or compensatory dynamics of the community's constituent species; and the potential for fluctuations and trends in species phenology to result in trophic mismatches. We find that positively associated fluctuations in the population densities of co-located species commonly show asymmetric tail associations; that is, it is common for two species' densities to be more correlated when large than when small, or vice versa. Ordinary measures of association such as correlation do not take this asymmetry into account. Likewise, positively associated fluctuations in the phenology of co-located species also commonly show asymmetric tail associations. We provide evidence that tail associations between two or more species' population-density or phenology time series can be inherited from mutual tail associations of these quantities with an environmental driver. We argue that our understanding of community dynamics and stability, and of phenologies of interacting species, can be meaningfully improved in future work by taking into account tail associations.

## KEYWORDS

aphids, copula, interspecies synchrony, match–mismatch hypothesis, plankton, tail association

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

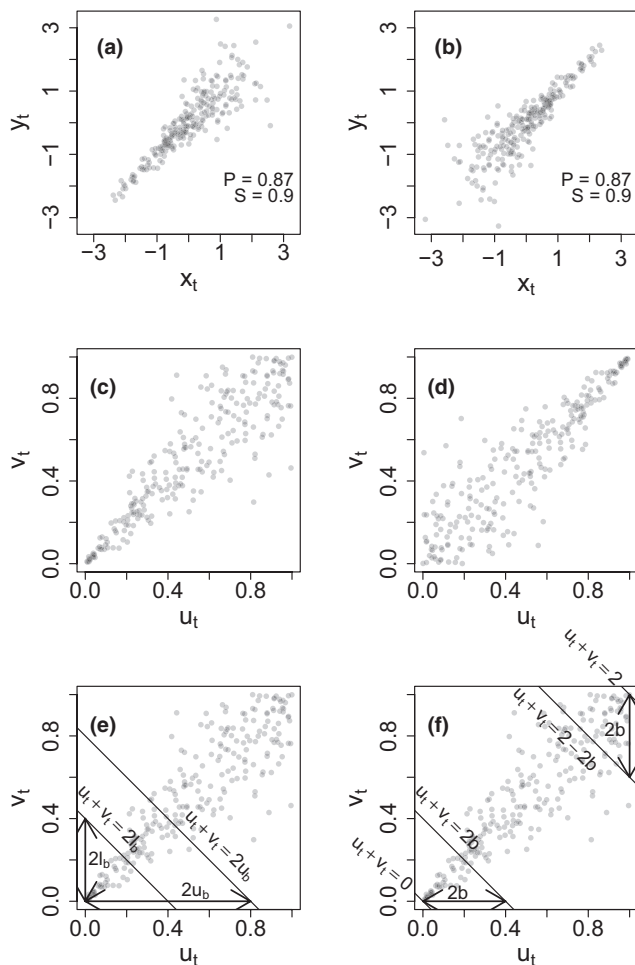
All ecologists study relationships between biological and environmental variables and among biological variables. But standard methods for studying the association between two variables provide only a small slice of the information content of the association. For instance, the two pairs of variables in Figure 1a,b have identical Pearson correlation coefficients, and also have identical Spearman correlation coefficients, but nonetheless display very different patterns of association (Ghosh, Sheppard, Holder, et al., 2020; Ghosh, Sheppard, & Reuman, 2020). Correlations are not the only way to study associations, but they are very commonly used, and other standard methods in ecology provide a similarly limited amount of information that neglects patterns of association (Anderson, de

Valpine, Punnett, & Miller, 2018; Genest & Favre, 2007; Joe, 2014; Mai & Scherer, 2017; Nelsen, 2006) that seem likely to be ecologically important (Ghosh, Sheppard, Holder, et al., 2020; Ghosh, Sheppard, & Reuman, 2020).

The variables of Figure 1a (respectively, Figure 1b) are more strongly related in the left (respectively, right) portions of their distributions, thereby displaying asymmetric associations of the distribution tails, henceforth called asymmetric *tail association*. For two positively associated variables, stronger association between values in the left or lower portions of the distributions of the variables is henceforth referred to as *left-tail association* (Figure 1a), whereas stronger association between values in the right or upper portions of the distributions of the variables is henceforth referred to as *right-tail association* (Figure 1b). The word “distribution” is sometimes omitted from the terminology, but implied. Tail association is a potentially important pattern of association that is not captured by standard correlation coefficients.

Statistical approaches exist, however, that provides a complete description of the relationship between variables; these approaches are based on the idea of the *copula*. Tail associations are an important aspect of a copula approach to dependence, and tail association will be a focus of this paper. We here give a conceptual flavor of copulas before subsequently focusing on tail association. We introduce copulas instead of proceeding directly to tail associations, for three reasons: to properly credit the copula ideas at the root of our tail association tools, and the researchers who developed them; to indicate the origin of our tail association tools, so that future researchers seeking to generalize our approach will have a place to start; and to introduce ideas (normalized rank plots—see below) that are necessary to define our measures of tail association. Copulas can be used to separate the information content of a bivariate dataset,  $(x_t, y_t)$  for  $t = 1, \dots, T$ , into two nonoverlapping parts: the information in the marginal distributions (which is not about the association between the variables) and the rest of the information (which is solely about the association). Following Ghosh, Sheppard, Holder, et al. (2020) and Genest and Favre (2007), the isolated information about the association between  $x_t$  and  $y_t$  is revealed by the plot of  $u_t$  against  $v_t$ , where  $u_t$  is the rank of  $x_t$  in the set  $\{x_1, x_2, \dots, x_T\}$ , divided by  $T + 1$ ; and  $v_t$  is the rank of  $y_t$  in the set  $\{y_1, y_2, \dots, y_T\}$ , also divided by  $T + 1$ . Here the rank of the smallest element of a set is understood to be 1. We refer to the  $u_t$  and  $v_t$  as *normalized ranks* of the  $x_t$  and  $y_t$ . We refer to the plot of  $v_t$  against  $u_t$  as the *normalized rank plot* for  $y_t$  and  $x_t$ . For instance, the normalized rank plots for Figure 1a,b are in Figure 1c,d and show the asymmetric associations in the tails. The normalized rank plot reflects the copula structure of  $(x_t, y_t)$  (Genest & Favre, 2007; Ghosh, Sheppard, Holder, et al., 2020). Ranking makes the marginal distributions uniform, isolating only the information on association between the variables. Genest and Favre (2007) states that inferences about dependence structures should always be based on ranks. It is likewise the purpose of copula approaches to separate association information from information on marginals.

We emphasize that we have not here provided a formal definition of copulas, instead only introducing the fundamental copula



**FIGURE 1** Pedagogical figure for introducing tail association and partial Spearman correlation. (a, b) Two pairs of variables that have identical Pearson ( $P$ ) correlation, and also identical Spearman ( $S$ ) correlation, but that differ markedly in the nature of the association. Panel a shows stronger left- than right-tail association and panel b shows the reverse. (c, d) Normalized rank plots (see Section 1) for panels a and b, respectively. (e, f) Graphics supporting the definitions of partial Spearman correlation and our statistic measuring asymmetry of tail association (see Section 2). This figure is similar in some respects to figs 1 and 7 of Ghosh, Sheppard, Holder, et al. (2020)

idea of separating dependence information from information on marginals. Brief (Anderson et al., 2018; Genest & Favre, 2007; Ghosh, Sheppard, Holder, et al., 2020) and comprehensive (Joe, 2014; Mai & Scherer, 2017; Nelsen, 2006) introductions to copulas are available elsewhere. Copulas can also be used to study multivariate data. Copula approaches are applied widely and to great effect in fields such as finance and neuroscience (Emura & Chen, 2016; Goswami, Hazra, & Goyal, 2018; Kim et al., 2008; Li, 2000; Li, Xie, & Hu, 2013; Onken, Grünwälder, Munk, & Obermayer, 2009; Serinaldi, 2008; She & Xia, 2018), but only rarely, so far, in ecology (Anderson et al., 2018; Ghosh, Sheppard, Holder, et al., 2020; Ghosh, Sheppard, & Reuman, 2020; Popovic, Warton, Thomson, Hui, & Moles, 2019; Valpine, Scranton, Knape, Ram, & Mills, 2014). The potential of copulas for improving ecological understanding was argued by Ghosh, Sheppard, Holder, et al. (2020), and those authors also introduced tail association as an important aspect of copula structure and elaborated the relationship between tail association and copulas.

The study of Ghosh, Sheppard, Holder, et al. (2020) was a wide-ranging study of the importance, causes, and consequences of copula structures in associations between ecological variables. One of the main foci of that paper was associations between fluctuations through time of population-density or phenological measurements of the same species in different locations. This study instead focuses on population-density and phenological measurements of different species in the same location. Ghosh, Sheppard, Holder, et al. (2020) studied, for instance, associations between first flight time series, for a given species of aphid, measured at different locations in the United Kingdom (UK); and associations between plankton density time series, for a given plankton taxon, measured at different locations in seas around the UK. We instead study associations between first flight or population-density time series measured in the same location for different (sympatric) species. Thus, in contrast with the study of Ghosh, Sheppard, Holder, et al. (2020), this study is more part of community ecology than of spatial ecology. Our reasons for this shift are as follows.

First, *synchronous* (positively correlated) and *compensatory* (negatively correlated) population-density dynamics of different species occupying the same area are longstanding topics of concern in community ecology, with important ramifications for the stability or instability of aggregate community or ecosystem properties (Gonzalez & Loreau, 2009; Jochimsen, Kümmerlin, & Straile, 2013; Kent, Yannarell, Rusak, Triplett, & McMahon, 2007; Loreau & Mazancourt, 2008; Raimondo, Turcáni, Patočka, & Liebhold, 2004); there are reasons to believe tail associations in this context will play an important but unstudied role in understanding these topics. A major past insight into community dynamics (Gonzalez & Loreau, 2009) was that an aggregate property of a community, such as its total biomass, can be relatively stable through time although its constituent parts (population biomasses of individual species) are highly variable, if the parts show compensatory dynamics (Hallett et al., 2014). Likewise, synchrony amplifies community biomass variability because the concordant variations of species biomass time

series reinforce each other in the total (Ma et al., 2017). If synchronous fluctuations show right-tail association, then species are highly abundant simultaneously, which may produce years of extremely high community biomass. Alternatively, if synchronous fluctuations show left-tail association, species are very scarce simultaneously, potentially producing years of extremely low community biomass. Thus the tail association of synchrony, not just the presence and strength of synchrony, may independently influence temporal variability of aggregate community properties. This is revisited in the Discussion.

Second, studies of the phenology of species interacting in one area have also played a central role in community ecology, with important ramifications for whether and to what extent interactions will be modified by climate change (Durant, Hjermmann, Ottersen, & Stenseth, 2007; Yang & Rudolf, 2010); there are reasons to believe tail associations between variables in this context may play an important role, as well. As climate changes and phenologies shift, there is the potential for phenologies of interacting species to shift differently, disrupting the interaction (Thackeray et al., 2010). This idea is referred to as the match–mismatch hypothesis. Even if, for instance, year-to-year fluctuations in the emergence times of two interacting species are highly correlated, if this correlation is principally in the right (respectively, left) tails of the distributions of possible emergence times, so that early (respectively, late) emergences of the species are actually uncorrelated, then mismatched years are likely to occur, impacting the species. Such mismatches will occur, in this conceptual example, when emergence is early (respectively, late). Essentially, even with substantial correlation between emergence dates of species, if this correlation is principally in one of the tails, then uncorrelated emergences, and therefore mismatches, can occur under some conditions. One potential mechanism by which early emergences, for example, may be uncorrelated between species while later emergences remain correlated is if both species follow the same environmental cue for their emergence, but physiological limitations of only one of the species prevent emergence before a certain date. Advancing emergence dates of myriad species make this scenario more plausible.

We here begin exploring whether tail associations may be important for studies of synchrony and compensatory dynamics, and for studies of phenology and the match–mismatch hypothesis. We use a 56-year dataset of population densities of 4 species of dinoflagellates from the *Ceratium* genus, from 15 locations in the seas around the UK; and a 35-year dataset of annual first flight dates for 20 species of aphid from 10 locations within the UK. The terms left- and right-tail association, defined above, do not apply to negatively associated variables, because the negative association means values in the left tail of one variable are associated with those in the right tail of the other; slightly modified methods are required to study tail association and its asymmetry in negatively associated variables. But our aphid and plankton population and phenology variables were almost exclusively positively associated with each other (see Section 3). Therefore, we introduce methods and present results in this study chiefly for the case of positively

associated variables, returning to the topics of negatively associated variables and compensatory population dynamics in the Discussion.

In addition to examining whether tail association in our data is asymmetric, we also test for possible causes of such patterns. One possible mechanism, similar to some of the mechanisms explored by Ghosh, Sheppard, Holder, et al. (2020), is explained for the *Ceratium* example as follows. Earlier work showed that average sea surface temperature is an important correlate of phytoplankton abundance in our data (e.g., Defriez, Sheppard, Reid, & Reuman, 2016; Sheppard, Defriez, Reid, & Reuman, 2019a; Sheppard, Reid, & Reuman, 2017): cold water is associated with more phytoplankton, likely because upwelling and mixing of the surface and deeper ocean layers bring both nutrients and cold water to the photic zone. However, if it is the case for a given location that very cold water is associated with no more *Ceratium*, on average, than is moderately cold water, then that corresponds to a positive relationship and a left-tail association between the "coldness" of the surface water (measured, for instance, by how many degrees colder the water is than average) and *Ceratium* abundance. If such tail association is strong and consistent across *Ceratium* species, it should produce positive relationships with left-tail association between the abundance time series of the species. Likewise, in locations for which the winter coldness-*Ceratium* abundance association shows less left-tail association, one should see less left-tail association between different *Ceratium* species. So tail association between two species may be inherited from joint tail association of both species on a common environmental driver. Phytoplankton are also strongly influenced by the abundant generalist copepod consumer *Calanus finmarchicus*, so our actual investigation of the mechanism proposed here will take into account this influence as well as the association with sea surface temperature. For aphid first flight, we examine the same potential mechanism, but the relevant driver in that case is winter temperature.

Thus this paper focuses on whether and why population-density or phenological time series of co-located species may show asymmetric patterns in their tail associations, with a focus on positively associated variables because positive associations are what occurred in the available data. We ask the following specific questions. (Q1) Do synchronous/positively correlated population-density or phenological time series of co-located species commonly show asymmetric tail associations? (Q2) If so, what are the causes of these patterns? We examine potential ecological consequences of asymmetric tail associations in the Discussion. We regard our investigation as a first step toward a better understanding of the potential importance of asymmetric tail associations for such central ecological topics as synchrony and compensatory dynamics in communities and their influence on community stability; and the match-mismatch hypothesis in phenology. The Discussion also has additional thoughts on next steps toward this goal. Our results and the conceptual considerations introduced above are good evidence, in our view, of the potential for tail association to make a

crucial difference in how ecologists understand these important topics.

## 2 | METHODS

### 2.1 | Data

Our population dataset comprised average annual abundance estimates for 15 locations (Figure S1) in the North Sea and British seas for 4 species from the *Ceratium* genus of dinoflagellates, and for the generalist consumer copepod species *C. finmarchicus*, for the 56 years 1958–2013. These data were a subset of a larger dataset covering 22 taxa and 26 locations, analyzed by Sheppard et al. (2017), Sheppard et al. (2019a), and Ghosh, Sheppard, Holder, et al. (2020). The locations are 2° by 2° grid cells. The data were originally obtained from the Continuous Plankton Recorder (CPR) dataset, now operated and maintained by the Marine Biological Association of the United Kingdom. Data preprocessing steps were the same as used by Ghosh, Sheppard, Holder, et al. (2020). *Ceratium* species were extracted in part because they have a role in harmful algal blooms (red tides) (Baek, Shimode, Shin, Han, & Kikuchi, 2009); and also because four species were available from the genus (Table 1), and we chose closely related species because they may be influenced in similar ways by environmental variables. The 15 locations we used were selected from the 26 locations of the larger dataset (Figure S1) as follows. First, to reduce the effects of sampling variation on statistical results, we chose the subset of locations for which more than 35 years of data were available for all species. Second, for a given location, we excluded *Ceratium* species that were undetected for more than 10% of sampled years at that location. Finally, we considered only those locations for which at least two *Ceratium* species remained. We also had data on average growing season sea surface temperature for each grid cell and year (Sheppard et al., 2017, 2019a). Earlier analyses (e.g., Sheppard et al., 2019a) demonstrated that sea surface temperature and *C. finmarchicus* abundance are important covariates of phytoplankton dynamics in UK seas, though associations between temperature and phytoplankton are probably due to relationships both these variables have with nutrient abundance in surface ocean layers. Sea surface temperature data preprocessing was the same as used by Sheppard et al. (2017).

Our phenology dataset comprised annual first flight dates for 20 aphid species (Table 1) from 10 locations across the UK (Figure S2), spanning the 35 years 1976–2010. These data were a subset of a larger dataset covering 11 locations, analyzed previously by Sheppard, Bell, Harrington, and Reuman (2016) and Ghosh, Sheppard, Holder, et al. (2020). The data were originally obtained from the Rothamsted Insect Survey suction-trap dataset (Bell et al., 2015; Harrington, 2014). Data preprocessing was the same as that of Sheppard et al. (2016). Locations were screened, leading to the removal of one of the original 11 sampling locations, by requiring at least 30 years of data be available for all species, again to reduce sampling variation of statistics. We also had time series of

**TABLE 1** Names of 4 plankton and 20 aphid species for which data were used

Species ID	Latin binomial
Plankton	
1	<i>Ceratium fusus</i>
2	<i>Ceratium furca</i>
3	<i>Ceratium tripos</i>
4	<i>Ceratium macroceros</i>
Aphids	
1	<i>Rhopalosiphum insertum</i>
2	<i>Rhopalosiphum padi</i>
3	<i>Aphis fabae</i>
4	<i>Sitobion fragariae</i>
5	<i>Hyperomyzus lactucae</i>
6	<i>Rhopalosiphum maidis</i>
7	<i>Nasonovia ribisnigri</i>
8	<i>Phorodon humuli</i>
9	<i>Sitobion avenae</i>
10	<i>Elatobium abietinum</i>
11	<i>Brachycaudus helichrysi</i>
12	<i>Brevicoryne brassicae</i>
13	<i>Hyalopterus pruni</i>
14	<i>Acyrtosiphon pisum</i>
15	<i>Myzus persicae</i>
16	<i>Macrosiphum euphorbiae</i>
17	<i>Metopolophium dirhodum</i>
18	<i>Myzus ascalonicus</i>
19	<i>Drepanosiphum platanoidis</i>
20	<i>Cavariella aegopodii</i>

winter average temperature for each location and year. The winter temperature for year  $t$  was the average of December of year  $t - 1$  to March of year  $t$ . Earlier analyses have demonstrated the importance of winter temperature for aphid first flight date (e.g., Sheppard et al., 2016).

## 2.2 | Statistical methods

Given bivariate data  $(x_t, y_t)$  for a set of years,  $t$ , of size  $T$ , and after computing normalized ranks  $(u_t, v_t)$  as described in the Introduction, tail association and asymmetry of tail association were measured using the *partial Spearman correlation* of Ghosh, Sheppard, Holder, et al. (2020), which we here reintroduce. The standard Spearman correlation itself measures association between the variables  $x_t$  and  $y_t$  (or between  $u_t$  and  $v_t$  – recall the Spearman correlation is based on ranks, so is the same for both sets of variables); but Spearman correlation measures only the overall association of the samples and cannot tell us how association varies across the distributions of the

variables. Given two bounds  $1 \leq l_b < u_b \leq 1$ , we define the boundary lines  $u + v = 2l_b$  and  $u + v = 2u_b$  (Figure 1e), which intersect the unit square on which normalized ranks are plotted. The partial Spearman correlation associated with the bounds  $l_b$  and  $u_b$  will be the portion of the Spearman correlation attributable to the points that fall between these boundary lines. The partial Spearman correlation for the band between these boundaries and within the unit square is,

$$\text{cor}_{l_b, u_b}(u, v) = \frac{\sum (u_t - \text{mean}(u))(v_t - \text{mean}(v))}{(T - 1) \sqrt{\text{var}(u)\text{var}(v)}} \quad (1)$$

Here, sample means and sample variances are computed using all  $T$  data points, but the sum,  $\sum$ , is over only the indices  $t$  for which  $u_t + v_t > 2l_b$  and  $u_t + v_t < 2u_b$ . The partial Spearman correlation is not defined if there are no points in the band. For positively associated  $(u_t, v_t)$ , the partial Spearman correlations  $\text{cor}_{0,b}$  and  $\text{cor}_{1-b,1}$  for  $b \leq 0.5$  (Figure 1f) measure association in the left and right tails, respectively, and can be compared via a difference,  $\text{cor}_{0,b} - \text{cor}_{1-b,1}$ , to measure asymmetry of tail association. Positive values (respectively, negative) of this difference mean stronger left-tail (respectively, right-tail) association. The sum of  $\text{cor}_{0,0.5}$  and  $\text{cor}_{0.5,1}$  (or the sum of  $\text{cor}_{l_b, u_b}$  for any other choice of bands  $(l_b, u_b)$  that partition  $(0, 1)$ ) equals the standard Spearman correlation, as long as no points happen to lie exactly on the bounds. Notation is summarized in Table S1.

For each sampling location,  $n$ , we computed a matrix,  $C^n$ , which we call the *community tail association matrix*, which quantifies asymmetry of tail association between pairs of aphid species or pairs of *Ceratium* species at  $n$ . Denote by  $s_i^n(t)$  the aphid first flight date or the *Ceratium* population-density for sampling location  $n$ , for the  $i$ th species that was present in the cleaned data for location  $n$ , and for year  $t$ . We then defined the matrix  $C^n$  by defining  $C^n(i, j)$  for two aphid or *Ceratium* species  $i, j$ , as follows. First,  $C^n(i, j)$  was not defined, or was defined to equal the missing-data space holder “NA”, if one of three conditions held true: (a)  $i = j$ ; or if (b) the hypothesis that  $s_i^n(t)$  and  $s_j^n(t)$  were independent could not be rejected (5% level, using a test described by Genest and Favre (2007), implemented in the function BiCopIndTest in the VineCopula package in R); or if (c) independence was rejected but the Spearman correlation of  $s_i^n(t)$  and  $s_j^n(t)$  was negative. Otherwise we defined  $C^n(i, j) = \text{cor}_{0,b}(s_i^n(t), s_j^n(t)) - \text{cor}_{1-b,1}(s_i^n(t), s_j^n(t))$ , where the partial Spearman correlations in this expression were computed over the times,  $t$ , for which data were available for location  $n$ . The entry  $C^n(i, j)$  was set to NA if independence of  $s_i^n(t)$  and  $s_j^n(t)$  could not be rejected because attempting to quantify tail association (or anything else about association) for independent variables is pointless.  $C^n(i, j)$  was set to NA for negatively associated  $s_i^n(t)$  and  $s_j^n(t)$  because negative association occurred for only one pair of species in one location in our data (plankton sampling location 18, species *C. furca* and *C. macroceros*, see Section 3). Tail association for negatively associated variables should be studied, and this topic is revisited in the Discussion, but negative associations were too rare in our data to study them. The community tail association matrix  $C^n$  is symmetric. The value  $b = 1/3$  was used for plankton



locations, whereas  $b=1/2$  was used for aphid locations because aphid time series were shorter, and larger  $b$  reduces sampling variation for our statistics (Ghosh, Sheppard, Holder, et al., 2020). See Appendix S1 for more information on the choice of  $b$ .

We also computed a matrix  $D^n$ , which we call the *community-driver tail association matrix*, which quantifies tail association between aphid or plankton time series and their covariates. Denote by  $d_k^n(t)$  the value of the  $k$ th covariate that operated at sampling location  $n$  in year  $t$  (winter temperature for an aphid sampling location, sea surface temperature or *C. finmarchicus* density for a *Ceratium* location). We then defined  $D^n$  by defining  $D^n(i, k)$  for an aphid or *Ceratium* species  $i$  and a covariate  $k$ , as follows. First,  $D^n(i, k)$  was not defined, or was set to NA, if the hypothesis that  $s_i^n(t)$  and  $d_k^n(t)$  were independent could not be rejected (5% level, BiCopIndTest). Otherwise, we either: (a) set  $D^n(i, k) = \text{cor}_{0,b}(s_i^n(t), d_k^n(t)) - \text{cor}_{1-b,1}(s_i^n(t), d_k^n(t))$  if  $s_i^n(t)$  and  $d_k^n(t)$  were positively associated (positive Spearman correlation); or (b) set  $D^n(i, k) = \text{cor}_{0,b}(s_i^n(t), -d_k^n(t)) - \text{cor}_{1-b,1}(s_i^n(t), -d_k^n(t))$  if  $s_i^n(t)$  and  $d_k^n(t)$  were negatively associated (negative Spearman correlation). For aphid first flight time series, for which  $k$  was always 1 and  $d_k^n(t)$  was winter temperature in location  $n$ , associations between  $s_i^n(t)$  and  $d_k^n(t)$  were always negative when they were significant (see Section 3). The same was true for *Ceratium* density time series and sea surface temperature. Thus our practice of using  $-d_k^n(t)$  was equivalent, in the case of temperature variables, to using a "coldness" index such as the number of degrees colder than an average or typical reference temperature, in place of temperature. Aphid and *Ceratium* data were always positively associated with the coldness index when they were significantly associated with it. Although *C. finmarchicus* abundance was positively associated with *Ceratium* time series in some sampling locations and negatively associated in others, it always showed the same sign of association with all *Ceratium* species within a location. Using  $-d_k^n(t)$  in place of  $d_k^n(t)$  when negative associations with aphid or *Ceratium* data occurred allowed us to study asymmetry of tail association using methods developed with positively associated variables in mind. We again used  $b=1/3$  for plankton data and covariates, and  $b=1/2$  for aphid data and winter temperature. For display, we horizontally concatenated the matrices  $C^n$  and  $D^n$  and displayed matrix values using color.

We used the community tail association matrix  $C^n$  for each sampling location  $n$  to answer Q1 from the Introduction, as follows. First, we counted the number,  $N_L^n$ , of entries of  $C^n$  which were not NA and which were greater than 0. These were the "left-tail dominant" species pairs, that is, pairs of species for which association was stronger in the left rather than in the right tails of the species distributions. We also counted the number,  $N_R^n$ , of right-tail dominant pairs, for which the corresponding entries of  $C^n$  were negative. If  $N_L^n$  was substantially greater than (respectively, substantially less than)  $N_R^n$  for a location  $n$ , it suggested that left-tail association (respectively, right-tail association) between species in that location was dominant, answering Q1 in the affirmative. We also calculated  $A_{C,L}^n$ , the sum of all positive, non-NA entries of  $C^n$ ;  $A_{C,R}^n$ , the sum of all negative, non-NA

entries of  $C^n$ ; and  $A_C^n = A_{C,L}^n + A_{C,R}^n$ , a general measure of asymmetry of tail association in location  $n$ . We refer to  $A_C^n$  as the *total community tail association*. We additionally calculated the normalized quantities  $F_{C,L}^n = A_{C,L}^n / (A_{C,L}^n + |A_{C,R}^n|)$  and  $F_{C,R}^n = A_{C,R}^n / (A_{C,L}^n + |A_{C,R}^n|)$ . Because  $0 \leq F_{C,L}^n \leq 1$ ,  $0 \leq |F_{C,R}^n| \leq 1$ , and  $F_{C,L}^n + |F_{C,R}^n| = 1$ , the relative sizes of  $F_{C,L}^n$  and  $|F_{C,R}^n|$  indicate the relative dominance of left- and right-tail association between species at location  $n$ . Together, all these statistics provide an answer to Q1.

We used the community tail association matrix,  $C^n$ , and the community-driver tail association matrix,  $D^n$ , to answer Q2 from the Introduction for the *Ceratium* and aphid data, as follows. First, we calculated  $A_D^n$ , the sum of all non-NA entries of  $D^n$ . This was analogous to  $A_C^n$ , but calculated using the matrix  $D^n$  instead of the matrix  $C^n$ . We refer to  $A_D^n$  as the *total community-driver tail association*. We then examined whether the values  $A_C^n$  and  $A_D^n$  were correlated across locations,  $n$ . This tests the causal hypothesis in the Introduction because it tests whether *Ceratium* or aphid time series having stronger right-tail (respectively, left-tail) association with environmental covariates in a given location also had stronger right-tail (respectively, left-tail) association with each other at that location. Recall that an environmental covariate was reversed (its negative was used) when it was negatively associated with a *Ceratium* or aphid species, and that no covariate was ever significantly positively associated with some *Ceratium* or aphid species and significantly negatively associated with another such species in the same location (see Section 3).

We also answered Q2 for the aphid data as follows. Within a location,  $n$ , for each species,  $i$ , we computed the mean  $\alpha_C^n(i)$  of all non-NA entries  $C^n(i, j)$ , for  $j$  ranging across all species for which we had data. This quantity measures an average tail association of species  $i$  with other species in the same location, with positive values for greater left-tail association and negative ones for greater right-tail association. We refer to  $\alpha_C^n(i)$  as the *species-community tail association* for species  $i$ . We then defined  $\alpha_D^n(i)$  as the sum of all non-NA entries  $D^n(i, k)$ , for  $k$  ranging across all covariates for which we had data. We refer to this as the *species-driver tail association* for species  $i$ . For aphids we only had one covariate, winter temperature, so  $\alpha_D^n(i) = D^n(i, k)$  for  $k = 1$  corresponding to winter temperature. We provide the more general definition of  $\alpha_D^n(i)$  that applies when more covariates were available so the definition can also be considered (briefly, see below) for *Ceratium* data. We then examined, for each location,  $n$ , whether  $\alpha_C^n(i)$  and  $\alpha_D^n(i)$  were correlated across species,  $i$ . This tests the causal hypothesis in the Introduction because it tests whether aphid species which were more right-tail (respectively, left-tail) associated with environmental covariates (winter temperature) also had time series that were more right-tail (respectively, left-tail) associated with the time series of other species in the location. Recall that winter temperature was always negatively associated with aphid first flight when it was significantly associated (see Section 3), and negative temperature (a coldness index) was used in computing  $D^n(i, k)$ . Testing whether  $\alpha_C^n(i)$  and  $\alpha_D^n(i)$  were correlated across species,  $i$ , within a location,  $n$ , was not practical for *Ceratium*, because we only had data for at most four

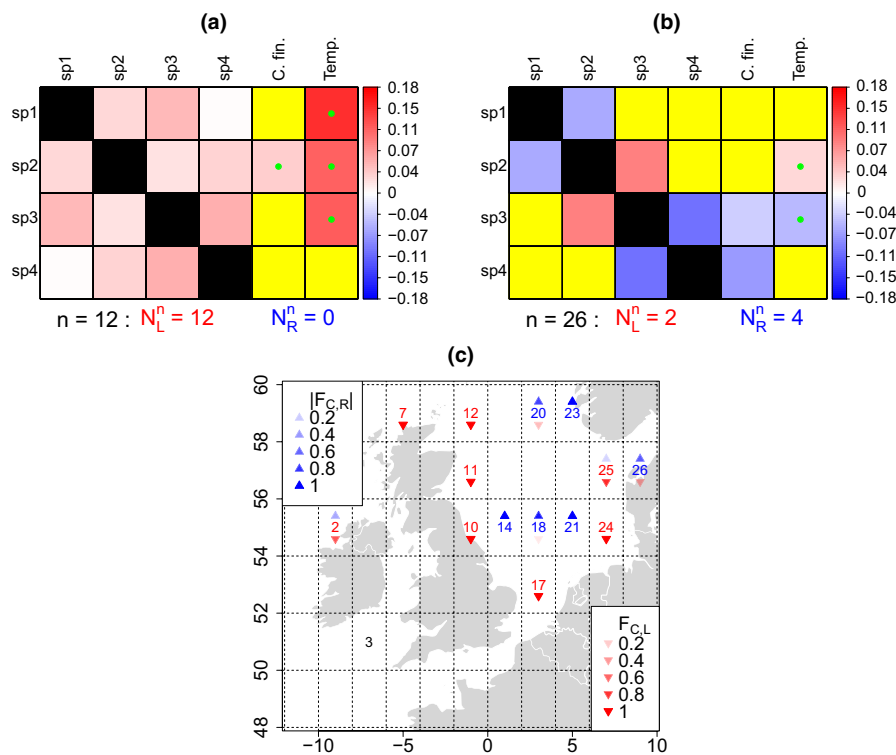
*Ceratium* species per sampling location, an insufficient number to provide much statistical power in testing for a correlation.

### 3 | RESULTS

Associations between *Ceratium* species were always positive when they were significant, except for one pair of species in one location (plankton sampling location 18, species *C. furca* and *C. macroceros*). Asymmetric tail association was very common between *Ceratium* population-density time series from the same location, answering Q1 in the affirmative for *Ceratium*; for some locations, left-tail association between *Ceratium* species was dominant, and for other locations right-tail association was dominant. To show this, we show that for some locations, the community tail association matrix,  $C^n$ , was comprised largely of positive values, indicating a preponderance of left-tail association between *Ceratium* time series for the location (Figure 2a). For such locations, *Ceratium* population densities are more likely to be correlated across species at low population

densities than at high densities. For other locations,  $C^n$  had mostly negative values, indicating a preponderance of right-tail association (Figure 2b). For such locations, *Ceratium* population densities are more likely to be correlated across species at high population densities than at low densities. To demonstrate the same result in another way, we show that the statistics  $F_{C,L}$  and  $F_{C,R}$  plotted across all sampling locations (Figure 2c), indicated that most *Ceratium* sampling locations were dominated by either left- or right-tail association, with approximately equal numbers of each, with only a few locations having more symmetric tail association, on average across pairs of *Ceratium* species.

Associations between aphid time series were always positive when they were significant. Asymmetric tail association was also very common between aphid first flight time series from the same location, answering Q1 in the affirmative for aphids; left-tail association was more common for some sampling locations and right-tail association dominated for others, but for most sites right-tail association dominated. To show this, we show that for some locations, the community tail association matrix,  $C^n$ , was comprised of a slight



**FIGURE 2** Either right- or left-tail association between population-density time series of *Ceratium* species could dominate, depending on the sampling location. (a, b) The community tail association matrix,  $C^n$ , and the community-driver tail association matrix,  $D^n$  (Statistical methods), horizontally concatenated, for example locations  $n = 12$  (a) and  $n = 26$  (b). See Table 1 for species names. All the non-NA values in  $C^n$  were positive (red) for location 12 (a), indicating left-tail association dominated in that location; but values were largely negative (blue) for location 26 (b), indicating right-tail association dominated there. Matrix entries which were NA because time series were independent are displayed in yellow. The counts  $N_L^n$  and  $N_R^n$  (see Section 2.2) also reflect the distinct tail association characteristics of the two locations. *C. fin.* = *C. finmarchicus*; Temp. = temperature. Green dots in  $D^n$  represent variables which were originally negatively associated, so the negative of the environmental covariate was used for calculating tail association. See Figure S3 for analogous figures for the other sampling locations. (c) The summary statistics  $F_{C,L}$  and  $F_{C,R}$  (see Section 2.2) for each site show that association between *Ceratium* species was either substantially dominated by the left or right tails of *Ceratium* distributions, with the exceptions of a few locations for which tail association was closer to symmetric. Site codes are colored red or blue depending on which of  $F_{C,L}$  or  $F_{C,R}$  had higher magnitude. Values are not plotted for site 3 because the hypothesis could not be rejected for that site that dynamics of distinct *Ceratium* species were independent

majority of positive values, indicating more left- than right-tail association between aphid time series for the location (Figure 3a); whereas for other locations,  $C^n$  had mostly negative values, indicating a preponderance of right-tail association (Figure 3b). To demonstrate the same result in another way, we show that the statistics  $F_{C,L}$  and  $F_{C,R}$  plotted across all sampling locations (Figure 3c), indicated that most aphid sampling locations had a preponderance of right-tail association, with only a few locations having more left-tail association, and those only slightly more. Thus, for most locations, aphid first flights are more correlated across species when first flights are later than average.

For the *Ceratium* data, the total community tail association,  $A_C^n$ , and the total community-driver tail association,  $A_D^n$  were significantly correlated across locations,  $n$ , validating our hypothesis from the Introduction for a cause of tail association between co-located species, and helping to answer Q2. In other words, tail association between co-located species time series was apparently inherited from common tail association of the species on environmental drivers. Across our 15 locations,  $A_C^n$  and  $A_D^n$  were significantly positively correlated (Pearson correlation, two-tailed test, Figure 4a). Thus locations for which *Ceratium* density time series showed greater left-tail (respectively, right-tail) association with environmental covariates (measured with  $A_D^n$ ) also exhibited greater left-tail (respectively, right-tail) association between density time series for distinct species (measured with  $A_C^n$ ).

For the aphid data, the total community tail association,  $A_C^n$ , and the total community-driver tail association,  $A_D^n$ , were positively but nonsignificantly correlated across our 10 sampling locations (Figure 4b). Thus locations for which aphid first flight time series showed greater left-tail (respectively, right-tail) association with winter temperature also showed a nonsignificant tendency toward greater left-tail (respectively, right-tail) association between the time series of distinct species. The correlation was close to significant for the aphid data, and may have been nonsignificant simply because there were slightly fewer aphid sampling locations than there were plankton locations. See also the subsequent results for aphids, which were significant and which support the same overall conclusions.

Our second analysis using aphids, based on the species-community tail associations,  $\alpha_C^n(i)$ , and the species-driver tail associations,  $\alpha_D^n(i)$  (Statistical methods), provided further evidence supporting our hypothesis for a cause of tail association between co-located species (Introduction). For 8 of 10 sampling locations,  $\alpha_C^n(i)$  and  $\alpha_D^n(i)$  were significantly correlated across species,  $i$  (Figure 5). In other words, for 8 of 10 locations, aphid species with greater left-tail (respectively, right-tail) association with winter temperature also had greater left-tail (respectively, right-tail) association with other aphid species.

## 4 | DISCUSSION

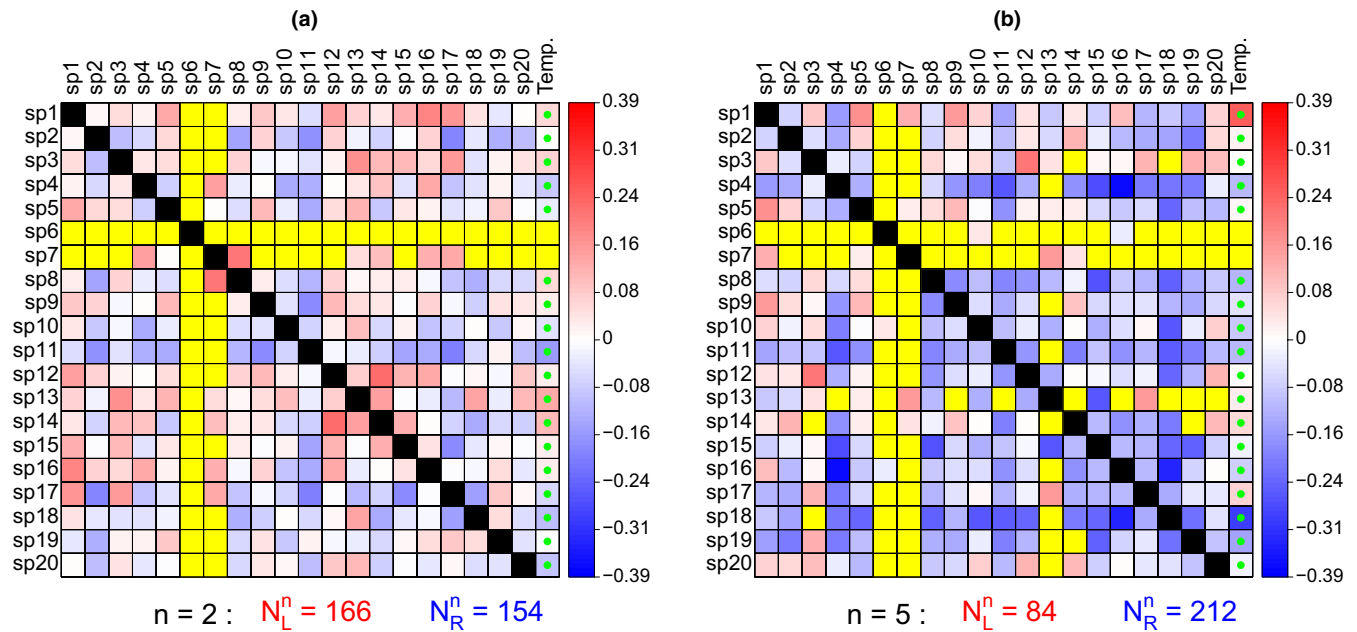
Our results show that synchronous population-density or phenological time series of co-located species can very commonly show

asymmetric tail association. For some sampling locations and species, tail association was predominantly in the left tails, and for others it was predominantly in the right tails of time series distributions, showing a new kind of ecologically meaningful variation among ecosystems. The partial Spearman correlation presented by Ghosh, Sheppard, Holder, et al. (2020) is a simple and effective way to measure tail association for ecological applications. Our results also demonstrate a mechanism by which asymmetric tail association between species can arise: It can be inherited by joint tail association of the two species on the same environmental variables. This mechanism seems likely to apply commonly when co-located species are influenced by the same external factors. Our results convincingly show that standard correlation approaches omit phenomena that seem likely to be important for at least two major topics of interest in ecology: synchronous/compensatory dynamics of species within a community and their influence on community stability; and shifting phenologies and the match–mismatch hypothesis.

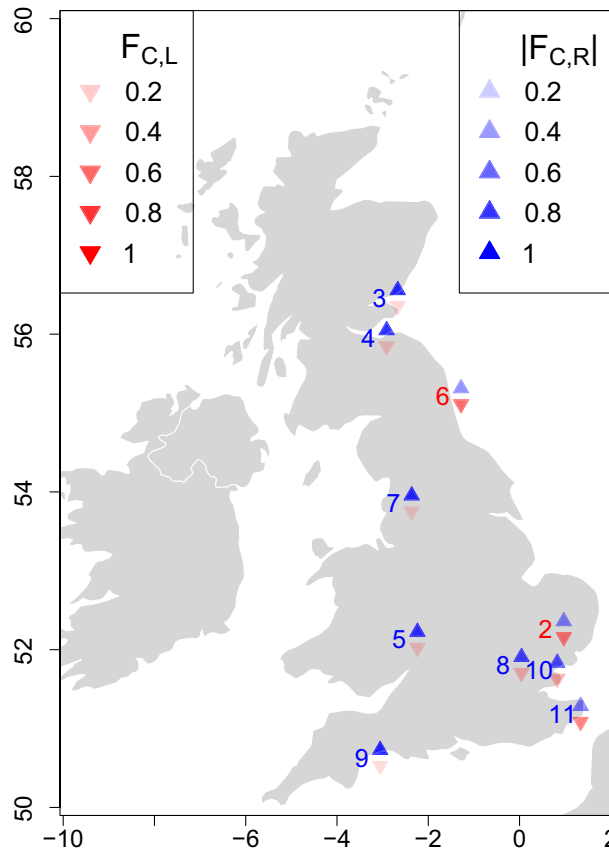
The distinct tail association characteristics of *Ceratium* in different sampling areas around the UK may have consequences for the stability through time of total *Ceratium* abundance, which may relate to harmful algal blooms because *Ceratium* species can have a role in such blooms (Baek et al., 2009). For locations in which left-tail association between *Ceratium* density time series is dominant, *Ceratium* species are scarce simultaneously, potentially producing years of very low total *Ceratium* biomass. In contrast, for locations in which right-tail association is dominant, *Ceratium* species are highly abundant simultaneously, which may produce years of very high *Ceratium* biomass, which may sometimes correspond to harmful algal blooms. Our results show that the distinction between these two types of location relates to the tail association of *Ceratium* species with their environmental covariates, sea surface temperature and *C. finmarchicus* density. It may be useful to study in future work why some locations principally have left-tail association with these drivers and some principally have right-tail association.

First flight time series for populations of co-located aphid species were principally right-tail associated; that is, more strongly correlated when first flights were later in the season. Our results show this was probably because: cold winters delay aphid first flights, but warm winters do not lead to first flights that are any earlier, on average, than those following moderate winters, producing right-tail association between first flights and winter coldness across multiple species; this common association leads to right-tail association between aphids. Thus winter temperature fluctuations lead to temporally dispersed early but temporally coordinated late arrival times of aphid species on summer hosts (many of which are crops, for the species we studied), a fact that may have pest-control significance. Winter temperature is known to influence the first flight dates of virtually all the aphid species for which we had data (Sheppard et al., 2016). Overwintering aphids are sensitive to frost conditions, and so winters probably reduce early spring populations on winter hosts plants. This then lengthens the time required for populations to reach sufficient densities to stimulate the production of winged morphs for flight to summer host plants.



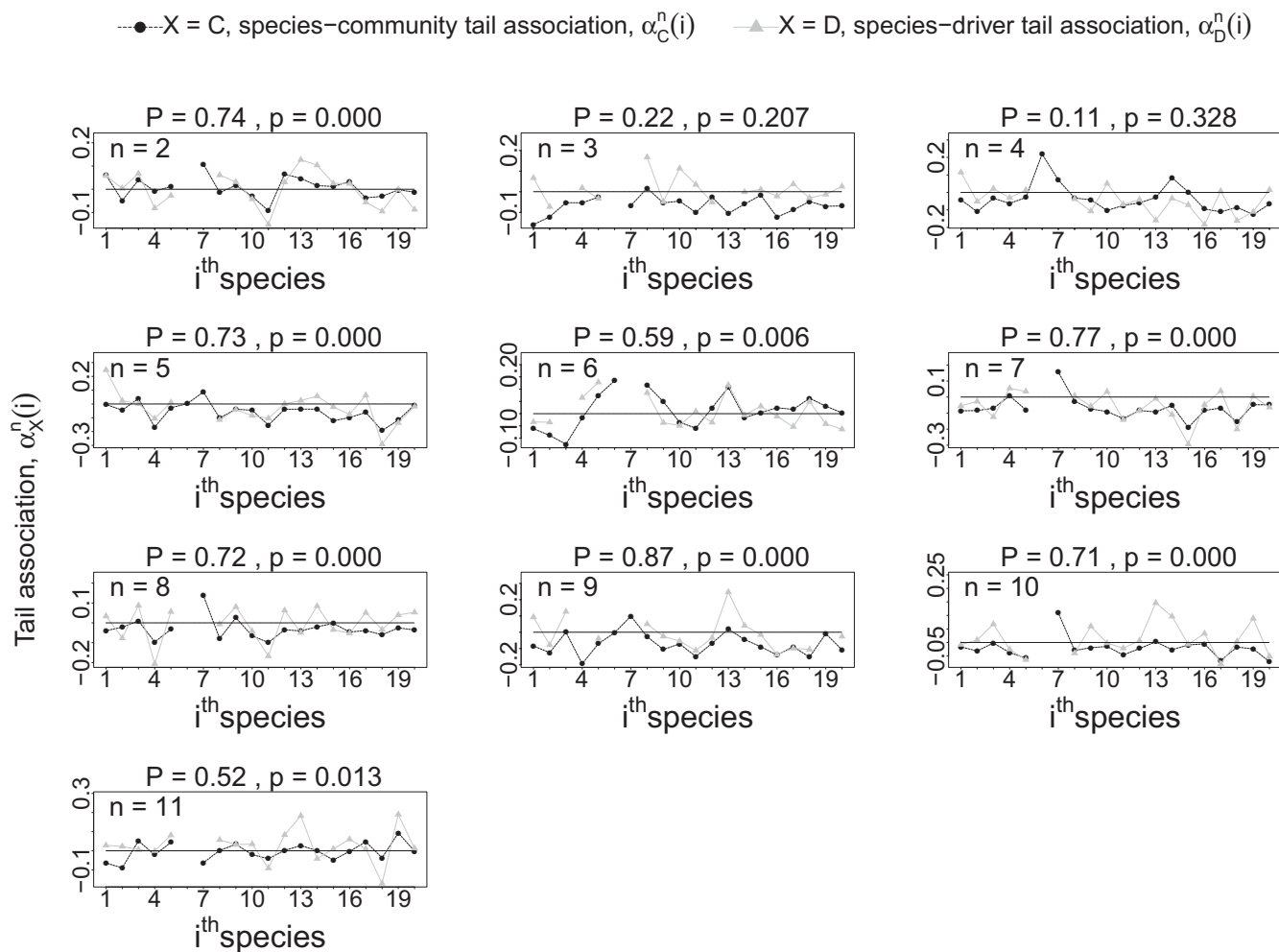
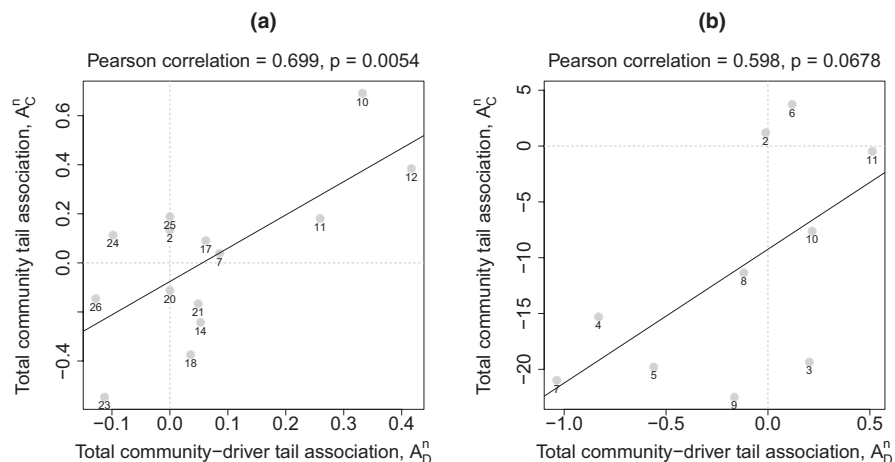


(c)



**FIGURE 3** Either right-tail association between first flight time series of aphid species could dominate, or left-tail association could be more common, depending on the sampling location. (a, b) The community tail association matrix,  $C^n$ , and the community-driver tail association matrix,  $D^n$  (Statistical methods), horizontally concatenated, for example locations  $n = 2$  (a) and  $n = 5$  (b). See Table 1 for species names. A slight majority of non-NA values in  $C^n$  were positive (red) for location 2 (a; see the  $N_L^n$  and  $N_R^n$  counts displayed), indicating left-tail association was slightly more common than right-tail association in that location. But values were largely negative (blue) for location 5 (b), indicating right-tail association dominated there. Matrix entries which were NA because time series were independent are displayed in yellow. Temp. = temperature. Green dots in  $D^n$  represent variables which were originally negatively associated, so the negative of winter temperature was used for calculating tail association (Statistical methods); this happened in all cases for which temperature and first flight were significantly associated. See Figure S4 for analogous figures for the other sampling locations. (c) The summary statistics  $F_{C,L}$  and  $F_{C,R}$  (see Section 2.2) for each site show that association was either dominated by the right tails, or, for a few locations, showed slightly more left-tail association. Site codes are colored red or blue depending on which of  $F_{C,L}$  or  $F_{C,R}$  had higher magnitude

**FIGURE 4** Tail association with environmental covariates was positively related to tail association between species for aphid and plankton time series. Panels show total community tail association,  $A_C^n$ , plotted against total community-driver tail association,  $A_D^n$  (Statistical methods), across locations,  $n$ , for *Ceratium* density (a) and aphid first flight (b) data. Pearson correlations and associated  $p$ -values for each panel are in the headers. Points are labeled with location numbers (see Figures S1 and S2)



**FIGURE 5** For 8 out of 10 sites, the Pearson correlation ( $P$ ) between the species–community tail association,  $\alpha_C^n(i)$ , and the species–driver tail association,  $\alpha_D^n(i)$ , across  $i = 1, 2, \dots, 20$ , was significantly positive ( $p < .05$ , one tailed test). This supports the hypothesis that tail association between species may be inherited from joint tail association of both species on a common environmental driver. See Table 1 for species IDs

If  $x_{s,l}(t)$  denotes the population-density of species  $s$  ( $s = 1, \dots, S$ ) in location  $l$  ( $l = 1, \dots, L$ ) at time  $t$ , we have here studied the nature and causes of tail association among the time series  $x_{s,l}(t)$  for a fixed  $l$  and for  $s = 1, \dots, S$ ; whereas Ghosh, Sheppard, Holder, et al. (2020) studied

the nature, causes and consequences of tail association among the time series  $x_{s,l}(t)$  for fixed  $s$  and  $l = 1, \dots, L$ , a distinct ecological context. One of the consequences studied by Ghosh, Sheppard, Holder, et al. (2020) relates to and illuminates a potential consequence,

mentioned above, of tail association for the ecological context of this study. Ghosh, Sheppard, Holder, et al. (2020) showed that the skewness, through time, of the spatial-total time series  $\sum_l x_{s,l}(t)$  is sensitive to the nature of tail association between the  $x_{s,l}(t)$  ( $l = 1, \dots, L$ ), if these time series are positively associated with each other. Right-tail (respectively, left-tail) association tended to produce right (respectively, left) skew in the total. Right skew corresponds to a spatial-total time series with exceptionally large values, that is, to “spiky”, unstable dynamics of the total population. Left skew corresponds to a spatial-total time series with low values, that is, to dynamics of the total population with a tendency to “crash”. The total population can be regarded as a landscape-level measure of the stability or variability of species  $s$ , and is important, for instance, if species  $s$  is a pest or an exploited species. For the same reasons, the skewness, through time, of the community-total time series  $\sum_s x_{s,l}(t)$  is sensitive to the tail association between the  $x_{s,l}(t)$  ( $s = 1, \dots, S$ ), which we have here studied. Right-tail (respectively, left-tail) association again tends to produce right (respectively, left) skew in the total time series. In this community context, the total is an aggregate property of the community, and the variability of this total has been used in an extensive literature (e.g., Hallett et al., 2014) to characterize community stability through time. This literature has explored the effects of synchronous versus compensatory dynamics in the  $x_{s,l}(t)$  ( $s = 1, \dots, S$ ) on the stability of the total community time series,  $\sum_s x_{s,l}(t)$ . But our results show that, even if all the species time series  $x_{s,l}(t)$  ( $s = 1, \dots, S$ ) are synchronous with each other, the tail association properties of these time series can influence the stability of the community-total time series.

Although our results are sufficient to show that tail associations are likely to be important for studies of community dynamics and stability, many communities show not only synchronous dynamics between some species pairs  $x_{s,l}(t)$  and  $x_{s',l}(t)$ , but also compensatory dynamics between other pairs. Our *Ceratium* time series were almost entirely synchronous, so we could not study the importance of tail association for compensatory dynamics. Next research steps should include the study of tail association between compensatory species within a local community. Furthermore, *Ceratium* is only part of the phytoplankton community in UK seas. It may be advantageous for future work to use data characterizing an entire competitive community. For instance, the data of Hallett et al. (2014) constitute annual abundances of all species of plant in an area. In that dataset, some species pairs show synchronous and some show compensatory dynamics.

Studying asymmetry of tail association for negatively correlated species density time series will require slightly modified methods. The only negative association between aphid or *Ceratium* time series that occurred in our system was not analyzed. Negative associations between species time series and the environmental covariates we considered were handled statistically by considering the positive association between the species time series and a “reversed” covariate; this corresponds to a positive association with a reconceptualized covariate, for example, a “coldness” index. But that approach would make no sense for negatively associated time series of two aphid or

*Ceratium* time series: there is no canonical choice of which variable to reverse. Asymmetry of tail association could still be considered, however, for negatively associated variables,  $u, v$ , in an unsigned approach, via the index  $|\text{cor}_{0,b}(u, 1-v) - \text{cor}_{1-b,1}(u, 1-v)|$ . Because  $|\text{cor}_{0,b}(u, 1-v) - \text{cor}_{1-b,1}(u, 1-v)| = |\text{cor}_{0,b}(1-u, v) - \text{cor}_{1-b,1}(1-u, v)|$  no choice need be made on which variable to “reverse.” A large value of this index indicates that tail association between  $u$  and  $v$  is asymmetric, though it does not provide information on whether association is stronger between the left tail of  $u$  and the right tail of  $v$  or between the right tail of  $u$  and the left tail of  $v$ .

Measures of tail association may also reveal useful information about freshwater plankton ecosystems and harmful algal blooms, in addition to information about marine harmful algal blooms (discussed above). Because blooms are extreme phenomena involving multiple species, monitoring the associations of phytoplankton species with each other and their associations with temperature and nutrient data in the extremes (this is tail association) could help us to better understand harmful blooms. Considering tail association may even produce improvements in statistics that have been developed to serve as early warning signals of impending major changes (so-called “tipping points”) in plankton communities and the lakes they inhabit (Buttitta, Carpenter, Loken, Pace, & Stanley, 2017; Carpenter et al., 2011), since some established early warning statistics make use of skewness of population distributions (Guttal & Jayaprakash, 2008). Tail association between phytoplankton species is related to skewness of the total phytoplankton biomass time series, as described in an earlier Discussion paragraph.

Although our aphid results were sufficient to demonstrate that tail association can be an important factor in the phenology of co-located species, it will be necessary in future work to apply tail association ideas to different datasets to assess whether these ideas can improve our understanding of the consequences of changing phenology for trophic phenological matching. The aphid species we studied have different host plants, so they do not directly interact. Shifts and fluctuations in the phenology of one species probably do not directly influence other species in our dataset. Future research should apply tail association to time series of phenologies of interacting species, such as the data on tree budburst dates, caterpillar abundance, and breeding phenology of great tits (*Parus major*) and blue tits (*P. caeruleus*) collected in Wytham Woods, Oxford, and other locations in Europe (e.g., Cole & Sheldon, 2017; Nilsson & Källander, 2006; Savill, Perrins, Kirby, & Fisher, 2011), or the extensive data collection from multiple trophic levels of Thackery et al. (2010).

One final idea for potentially valuable future research has to do with combining our approach, based on tail associations, with other recent approaches which emphasize other statistical aspects of the synchrony. For instance, research has now showed that synchrony and compensatory dynamics in communities have “timescale structure”; that is, the dynamics of two or more species can be synchronous on some timescales of analysis and compensatory on others (Keitt & Fisher, 2006; Vasseur et al., 2015; Zhao et al., 2020). How timescale specificity and tail associations interact is unknown, but potentially interesting. Multivariate copula approaches (Czado, 2019;

Joe, 2014) may be useful in this and other future extensions of the work we have begun here.

Our results extend the results of Ghosh, Sheppard, Holder, et al. (2020). Those authors argued that considering copulas and tail associations can provide insights across the field of ecology. But Ghosh, Sheppard, Holder, et al. (2020) did not consider co-located species, a context important for community ecology which we considered here.

## ACKNOWLEDGMENTS

We thank the many contributors to the large datasets we used; D. Stevens and P. Verrier for data extraction; and Joel E. Cohen, Lauren Hallett, and Jonathan Walter for helpful suggestions. We thank James Bell of the Rothamsted Insect Survey (RIS). The RIS, a UK Capability, is funded by the Biotechnology and Biological Sciences Research Council under the Core Capability Grant BBS/E/C/000J0200. SG, LWS and DCR were partly funded by US National Science Foundation grants 1714195 and 1442595 and the James S McDonnell Foundation.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTIONS

**Shyamolina Ghosh:** Conceptualization (equal); Formal analysis (lead); Methodology (equal); Software (lead); Visualization (lead); Writing-original draft (lead); Writing-review & editing (equal). **Lawrence W. Sheppard:** Conceptualization (equal); Formal analysis (supporting); Writing-review & editing (supporting). **Philip C. Reid:** Data curation (lead); Writing-review & editing (supporting). **Daniel C Reuman:** Conceptualization (equal); Formal analysis (supporting); Funding acquisition (lead); Methodology (equal); Supervision (lead); Writing-original draft (equal); Writing-review & editing (equal).

## DATA AVAILABILITY STATEMENT

Plankton data are available from the Dryad Digital Repository <https://doi.org/10.5061/dryad.rq3jc84> (Sheppard, Defriez, Reid, & Reuman, 2019b). Full code for the analyses can be downloaded from Dryad Digital Repository [https://datadryad.org/stash/share/eC20ojo\\_e9UTmXAoX1oq3klfT3aY0iptcUqAk8MHRAA](https://datadryad.org/stash/share/eC20ojo_e9UTmXAoX1oq3klfT3aY0iptcUqAk8MHRAA).

## ORCID

Shyamolina Ghosh  <https://orcid.org/0000-0002-5137-9933>

Lawrence W. Sheppard  <https://orcid.org/0000-0002-9032-2059>

Daniel Reuman  <https://orcid.org/0000-0002-1407-8947>

## REFERENCES

- Anderson, M., de Valpine, P., Punnett, A., & Miller, A. (2018). A pathway for multivariate analysis of ecological communities using copulas. *Ecology and Evolution*, 9, 3276–3294. <https://doi.org/10.1002/ece3.4948>
- Baek, S., Shimode, S., Shin, K., Han, M.-S., & Kikuchi, T. (2009). Growth of dinoflagellates, *Ceratium furca* and *Ceratium fusus* in Sagami Bay, Japan: The role of vertical migration and cell division. *Harmful Algae*, 8, 843–856. <https://doi.org/10.1016/j.hal.2009.04.001>
- Bell, J. R., Alderson, L., Izera, D., Kruger, T., Parker, S., Pickup, J., ... Harrington, R. (2015). Long-term phenological trends, species accumulation rates, aphid traits and climate: Five decades of change in migrating aphids. *Journal of Animal Ecology*, 84, 21–34. <https://doi.org/10.1111/1365-2656.12282>
- Butitta, V. L., Carpenter, S. R., Loken, L. C., Pace, M. L., & Stanley, E. H. (2017). Spatial early warning signals in a lake manipulation. *Ecosphere*, 8, e01941. <https://doi.org/10.1002/ecs2.1941>
- Carpenter, S. R., Cole, J. J., Pace, M. L., Batt, R., Brock, W. A., Cline, T., ... Weidel, B. (2011). Early warnings of regime shifts: A whole-ecosystem experiment. *Science*, 332, 1079–1082. <https://doi.org/10.1126/science.1203672>
- Cole, E., & Sheldon, B. (2017). The shifting phenological landscape: Within- and between-species variation in leaf emergence in a mixed-deciduous woodland. *Ecology and Evolution*, 7, 1135–1147. <https://doi.org/10.1002/ece3.2718>
- Czado, C. (2019). *Analyzing dependent data with vine copulas: A practical guide with R*. New York, NY: Springer.
- de Valpine, P., Scranton, K., Knape, J., Ram, K., & Mills, N. (2014). The importance of individual development variation in stage-structured population models. *Ecology Letters*, 17, 1026–1038.
- Defriez, E. J., Sheppard, L. W., Reid, P. C., & Reuman, D. C. (2016). Climate change-related regime shifts have altered spatial synchrony of plankton dynamics in the North Sea. *Global Change Biology*, 22, 2069–2080. <https://doi.org/10.1111/gcb.13229>
- Durant, J., Hjermann, D., Ottersen, G., & Stenseth, N. (2007). Climate and the match or mismatch between predator requirements and resource availability. *Climate Research*, 33, 271–283. <https://doi.org/10.3354/cr033271>
- Emura, T., & Chen, Y.-H. (2016). Gene selection for survival data under dependent censoring: A copula-based approach. *Statistical Methods in Medical Research*, 25, 2840–2857. <https://doi.org/10.1177/0962280214533378>
- Genest, C., & Favre, A.-C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, 12, 347–368. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2007\)12:4\(347\)](https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(347))
- Ghosh, S., Sheppard, L. W., Holder, M. T., Loecke, T. E., Reid, P. C., Bever, J. D., & Reuman, D. C. (2020). Copulas and their potential for ecology. *Advances in Ecological Research*, 62, 409–468.
- Ghosh, S., Sheppard, L. W., & Reuman, D. C. (2020). Tail associations in ecological variables and their impact on extinction risk. *Ecosphere*, 11, e03132. <https://doi.org/10.1002/ecs2.3132>
- Gonzalez, A., & Loreau, M. (2009). The causes and consequences of compensatory dynamics in ecological communities. *Annual Review of Ecology, Evolution, and Systematics*, 40, 393–414. <https://doi.org/10.1146/annurev.ecolsys.39.110707.173349>
- Goswami, U. P., Hazra, B., & Goyal, M. K. (2018). Copula-based probabilistic characterization of precipitation extremes over North Sikkim Himalaya. *Atmospheric Research*, 212, 273–284. <https://doi.org/10.1016/j.atmosres.2018.05.019>
- Guttal, V., & Jayaprakash, C. (2008). Changing skewness: An early warning signal of regime shifts in ecosystems. *Ecology Letters*, 11, 450–460. <https://doi.org/10.1111/j.1461-0248.2008.01160.x>
- Hallett, L. M., Hsu, J. S., Cleland, E. E., Collins, S. L., Dickson, T. L., Farrer, E. C., ... Suding, K. N. (2014). Biotic mechanisms of community stability shift along a precipitation gradient. *Ecology*, 95, 1693–1700. <https://doi.org/10.1890/13-0895.1>
- Harrington, R. (2014). The Rothamsted Insect Survey strikes gold. *Antenna*, 38, 159–166.

- Jochimsen, M. C., Kümmerlin, R., & Straile, D. (2013). Compensatory dynamics and the stability of phytoplankton biomass during four decades of eutrophication and oligotrophication. *Ecology Letters*, *16*, 81–89. <https://doi.org/10.1111/ele.12018>
- Joe, H. (2014). *Dependence modeling with copulas*. Chapman & hall/crc monographs on statistics & applied probability. New York, NY: Taylor & Francis.
- Keitt, T., & Fisher, J. (2006). Detection of scale-specific community dynamics using wavelets. *Ecology*, *87*, 2895–2904. [https://doi.org/10.1890/0012-9658\(2006\)87\[2895:DOSCDU\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[2895:DOSCDU]2.0.CO;2)
- Kent, A. D., Yannarell, A. C., Rusak, J. A., Triplett, E. W., & McMahon, K. D. (2007). Synchrony in aquatic microbial community dynamics. *The ISME Journal*, *1*, 38. <https://doi.org/10.1038/ismej.2007.6>
- Kim, J.-M., Jung, Y.-S., Sungur, E. A., Han, K.-H., Park, C., & Sohn, I. (2008). A copula method for modeling directional dependence of genes. *BMC Bioinformatics*, *9*, 225. <https://doi.org/10.1186/1471-2105-9-225>
- Li, D. (2000). On default correlation: A copula function approach. *Journal of Fixed Income*, *9*, 43–54. <https://doi.org/10.3905/jfi.2000.319253>
- Li, Y., Xie, K., & Hu, B. (2013). Copula-ARMA model for multivariate wind speed and its applications in reliability assessment of generating systems. *Journal of Electrical Engineering and Technology*, *8*, 421–427. <https://doi.org/10.5370/JEET.2013.8.3.421>
- Loreau, M., & de Mazancourt, C. (2008). Species synchrony and its drivers: Neutral and nonneutral community dynamics in fluctuating environments. *The American Naturalist*, *172*, E48–E66. <https://doi.org/10.1086/589746>
- Ma, Z., Liu, H., Mi, Z., Zhang, Z., Wang, Y., Xu, W., ... He, J.-S. (2017). Climate warming reduces the temporal stability of plant community biomass production. *Nature Communications*, *8*, 15378. <https://doi.org/10.1038/ncomms15378>
- Mai, J.-F., & Scherer, M. (2017). *Simulating copulas: Stochastic models, sampling algorithms, and applications*, 2nd ed. London, UK: Imperial College Press.
- Nelsen, R. B. (2006). *An introduction to copulas*, 2nd ed. New York, NY: Springer.
- Nilsson, J.-A., & Källander, H. (2006). Leafing phenology and timing of egg laying in great tits *Parsus major* and blue tits *P. caeruleus*. *Journal of Avian Biology*, *37*, 357–363.
- Onken, A., Grünwälder, S., Munk, M., & Obermayer, K. (2009). Analyzing short-term noise dependencies of spike-counts in macaque prefrontal cortex using copulas and the flashlight transformation. *PLoS Computational Biology*, *11*, e1000577. <https://doi.org/10.1371/journal.pcbi.1000577>
- Popovic, G. C., Warton, D. I., Thomson, F. J., Hui, F. K., & Moles, A. T. (2019). Untangling direct species associations from indirect mediator species effects with graphical models. *Methods in Ecology and Evolution*, *10*, 1571–1583. <https://doi.org/10.1111/2041-210X.13247>
- Raimondo, S., Turcáni, M., Patoëka, J., & Liebhold, A. M. (2004). Interspecific synchrony among foliage-feeding forest lepidoptera species and the potential role of generalist predators as synchronizing agents. *Oikos*, *107*, 462–470. <https://doi.org/10.1111/j.0030-1299.2004.13449.x>
- Savill, P., Perrins, C., Kirby, K., & Fisher, N. (2011). *Wytham Woods: Oxford's ecological laboratory*. New York, NY: Oxford University Press.
- Serinaldi, F. (2008). Analysis of inter-gauge dependence by Kendall's  $\tau_K$ , upper tail dependence coefficient, and 2-copulas with application to rainfall fields. *Stochastic Environmental Research and Risk Assessment*, *22*, 671–688. <https://doi.org/10.1007/s00477-007-0176-4>
- She, D., & Xia, J. (2018). Copulas-based drought characteristics analysis and risk assessment across the loess plateau of China. *Water Resources Management*, *32*, 547–564. <https://doi.org/10.1007/s11269-017-1826-z>
- Sheppard, L. W., Bell, J. R., Harrington, R., & Reuman, D. C. (2016). Changes in large-scale climate alter spatial synchrony of aphid pests. *Nature Climate Change*, *6*, 610. <https://doi.org/10.1038/nclimate2881>
- Sheppard, L. W., Defriez, E. J., Reid, P. C., & Reuman, D. C. (2019a). Synchrony is more than its top-down and climatic parts: Interacting Moran effects on phytoplankton in British seas. *PLoS Computational Biology*, *15*, e1006744. <https://doi.org/10.1371/journal.pcbi.1006744>
- Sheppard, L. W., Defriez, E. J., Reid, P. C., & Reuman, D. C. (2019b). Data from: Synchrony is more than its top-down and climatic parts: Interacting Moran effects on phytoplankton in British seas. *Dryad*. <https://doi.org/10.5061/dryad.rq3jc84>
- Sheppard, L. W., Reid, P. C., & Reuman, D. C. (2017). Rapid surrogate testing of wavelet coherences. *EPJ Nonlinear Biomedical Physics*, *5*, 1. <https://doi.org/10.1051/epjnbp/2017000>
- Thackeray, S., Sparks, T., Frederiksen, M., Burthe, S., Bacon, P., Bell, J., Botham, M. et al (2010). Trophic level asynchrony in rates of phenological changes for marine, freshwater and terrestrial environments. *Global Change Biology*, *16*, 3304–3313.
- Vasseur, D., Fox, J., Gonzalez, A., Adrian, R., Beisner, B., Helmus, M., ... Steiner, C. F. (2015). Synchronous dynamics of zooplankton competitors prevail in temperate lake ecosystems. *Proceedings of the Royal Society B*, *281*, 20140633.
- Yang, L., & Rudolf, V. (2010). Phenology, ontogeny and the effects of climate change on the timing of species interactions. *Ecology Letters*, *13*, 1–10. <https://doi.org/10.1111/j.1461-0248.2009.01402.x>
- Zhao, L., Wang, S., Hallett, L. M., Rypel, A. L., Sheppard, L. W., Castorani, M. C. N., ... Reuman, D. C. (2020). A new variance ratio metric to detect the timescale of compensatory dynamics. *Ecosphere*, *11*, e03114. <https://doi.org/10.1002/ecs2.3114>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Ghosh S, Sheppard LW, Reid PC, Reuman D. A new approach to interspecific synchrony in population ecology using tail association. *Ecol Evol* 2020;10:12764–12776. <https://doi.org/10.1002/ece3.6732>