Temporal origin of nestedness in interaction networks

Phillip P. A. Staniczenko,[1,*] Debabrata Panja[2,3]

[1]City University of New York (CUNY), Brooklyn College,
2900 Bedford Avenue, Brooklyn, NY 11210, USA
[2]Department of Information and Computing Sciences, Utrecht University,
Princetonplein 5, 3584 CC Utrecht, The Netherlands
[3]Centre for Complex Systems Studies, Utrecht University,
Minnaertgebouw, Leuvenlaan 4, 3584 CE Utrecht, The Netherlands
[*]To whom correspondence should be addressed; E-mail: pstaniczenko@brooklyn.cuny.edu

Abstract

Nestedness is a common property of communication, finance, trade, and ecological networks. In networks with high levels of nestedness, the link positions of low-degree nodes (those with few links) form nested subsets of the link positions of high-degree nodes (those with many links), leading to matrix representations with characteristic upper-triangular or staircase patterns. Recent theoretical work has connected nestedness to the functionality of complex systems and has suggested it is a structural by-product of the skewed degree distributions often seen in empirical data. However, mechanisms for generating nestedness remain poorly understood, limiting the connections that can be made between system processes and observed network structures. Here, we show that a simple probabilistic model based on phenology — the timing of co-presences among interaction partners — can produce nested structures and correctly predict around two-thirds of interactions in two fish market networks and around one-third of interactions in 22 plant-pollinator networks. Notably, the links most readily explained by frequent actor co-presences appear to form a backbone of nested interactions, with the remaining interactions attributable to opportunistic interactions or preferences for particular interaction partners that are not routinely available.

Significance statement: Networks describe the relationships among actors in complex systems. In nested networks, actors involved in few interactions are connected to actors involved in many interactions, with those highly-connected actors also interacting with other highly-connected actors. This pattern is seen in a variety of empirical systems and influences the response to external perturbations, but little is known about the processes that give rise to nestedness. We show that phenology, the day-to-day timing of interaction partner availability, is a general mechanism that generates nested structures. We present a simple probabilistic model which accounts for actor overlap through time but assumes actors have no preference for specific interaction partners, thereby providing an instructive baseline for investigating higher-level selection processes in interaction networks.

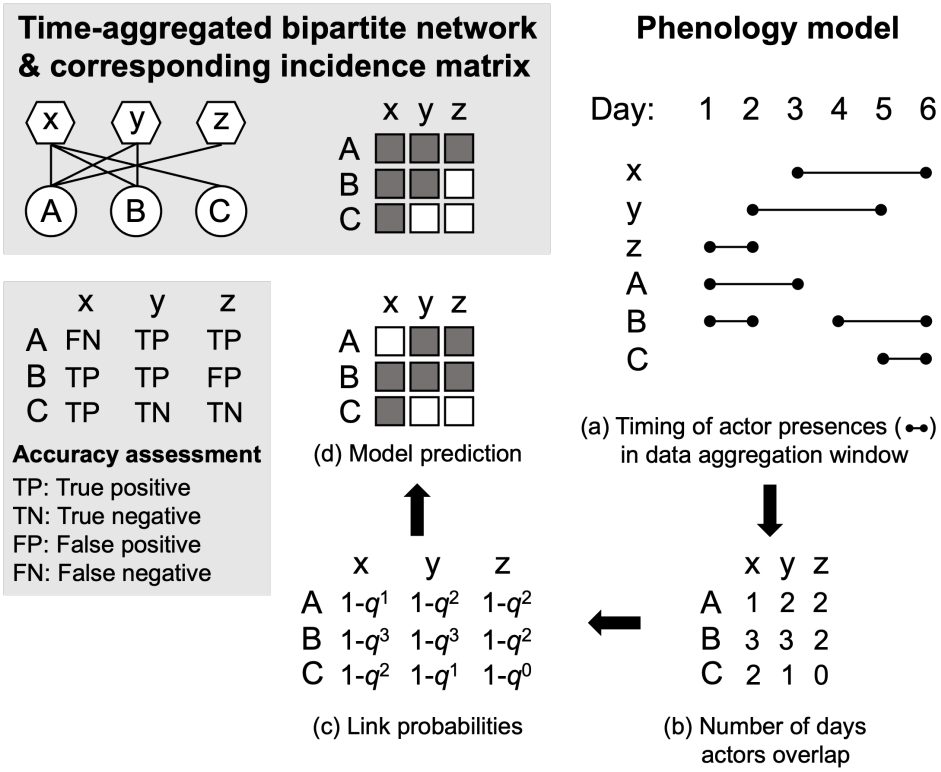

Figure 1: Example of a perfectly nested bipartite network (upper box), outline of phenology model, and model assessment scheme (lower box). A link is included in an empirical network if an interaction is recorded during the data aggregation window between two actors of different types, node sets {A,B,C} and {x,y,z}, representing, e.g., different plant and pollinator species, respectively. In the corresponding incidence matrix, filled entries indicate links between nodes (rows and columns) and nestedness is characterized by an upper-triangular or staircase pattern, as shown. The phenology model starts with a description of when actors are present and active in the system during the data aggregation window, which in this case is 6 days (a). This temporal information dictates the number of co-presences between each pair of nodes (b), which is used to determine the probability of each link in model-generated networks (c). Specifically, the probability of an interaction given a co-presence is assumed to be p and the probability of no interaction $q = 1 - p$, and so the probability of at least one interaction during the data aggregation window given n co-presences is $1 - q^n$. Link probabilities are scaled such that the expected number of links for an ensemble of model-generated networks equals the number in the empirical network (*Methods*). Each model realization returns an incidence matrix (d) and each link can assigned to one of four categories: true positive (TP) — link present in both model and empirical networks; true negative (TN) — link absent in both model and empirical networks; false positive (FP) — link present in model network but absent in empirical network; false negative (FN) — link absent in model network but present in empirical network.

The concept of nestedness first gained popularity in the field of island biogeography, where it was shown that the species found on distant islands formed progressive subsets of the species found on islands closer to the mainland [1]. Nestedness has since been studied in a wide variety of networks, most notably in bipartite networks in which links represent relationships between two different categories of actors [2] (an illustration of the characteristic pattern of nestedness is shown in the upper-left panel in Fig. 1). In some networks, links represent associations, e.g., product imports/exports between countries [3], while in others, links represent genuine interactions among system actors, e.g., visitation events between insect pollinators and flowering plants [4]. For interaction networks in particular, nestedness has been proposed to affect the feasibility [5], stability [6, 7], robustness [8], and persistence [9] of complex systems.

While the motivations for studying nestedness are grounded in well-established frameworks that provide general explanations for why nested structures are plausible — e.g., evolutionary theory and trait matching [10–12], auction theory and loyalty [8, 13] — less attention has been paid to uncovering the *mechanisms* that can generate nestedness. Some process-based models for generating realistic network structures have been proposed, especially in ecology [14–16], but the focus so far has been on the most appropriate ways of measuring nestedness [17] and, correspondingly, determining whether empirical networks are significantly nested [18]. Various null models have been developed and those that conserve observed degree distributions have been particularly successful at explaining the levels of nestedness seen in empirical networks [19, 20]. These findings have led to further work demonstrating that skewed degree distributions impose topological constraints that make nested structures likely [21], and therefore that the ubiquity of nestedness should be considered unsurprising.

One reason for the difficulty in identifying generative mechanisms is that nestedness has typically been studied in single networks built from time-aggregated data that mask processes occurring at time scales shorter than the aggregation window. However, recent efforts have led to the collection of highly time-

resolved interaction data [22, 23], which provides an opportunity to explore how the cyclic and seasonal dimensions of social and natural phenomena relate to nestedness at different temporal scales. In ecology, the study of periodic biological events is known as phenology, and various temporal processes have been proposed to explain the overall structure of mutualistic interaction networks [24–26]. Although current explanations for nestedness presuppose elaborate system-specific processes, e.g., trait matching between insect proboscis length and flower corolla depth formed over evolutionary time scales for plant-pollinator networks [4], the availability of time-resolved interaction data means that the effect of more general yet still fundamental processes at shorter time scales, such as how often potential interaction partners overlap through time, can now be investigated. We found that a simple probabilistic model based on phenology could explain and predict the temporal emergence of nested interaction patterns in a broad range of systems. Specifically, the model produced increasingly nested networks as additional temporal data were incorporated, ultimately generating empirically observed levels of nestedness when the amount of temporal aggregation matched a complete phenological cycle.

Results

Phenology model

We analyzed a total of 24 previously published data sets that are publicly available: two of buyer-seller transactions in a fish market in Boulogne-sur-Mer, France [22] and 22 of plant-pollinator visitations sampled across nine countries [23]. All data sets include interactions resolved to the daily level, with data spanning at least one year, or in the case of the ecological data, at least one complete flowering season within a single year (note that data were not collected for all days in an overall sampling period; *Methods*). For networks built from the fish market data, nodes represent individual buyers and sellers and links represent at least one transaction of money for goods. One of the two fish market data sets corresponds to an auction market, where buyers and sellers were matched anonymously by an order book; for the other, buyers and sellers could negotiate directly with one another to agree on prices. For the plant-pollinator data, nodes represent species and links represent the observation of at least one visit between individuals belonging to the respective species. To the best of our knowledge, these two groups (fish market and plant-pollinator) of data sets are the only freely available data sets with daily time-resolved interaction data.

For both the fish market and plant-pollinator data sets, interactions (links) could only be realized if system actors (nodes) were physically present together at the same place on a given day; if a particular actor (buyer, seller, pollinator, plant) was not involved in *any* interaction on a given day, then we assume the actor was not around that day. This simple yet important temporal feature — which is shared by most types of interaction networks — is rarely accounted for when analyzing networks built from aggregated data, despite clearly impacting which interactions are more or less likely to be recorded. To explore this effect on empirical network structures and their levels of nestedness, we developed the *phenology model*, which assumes the probability of observing a link increases with the number of times each pair of actors was co-present within a pre-defined aggregation window (Fig. 1 and *Methods*). As our main point of comparison, we used the *degree distribution model*, which conserves the degree distributions of empirical networks, since it is the current best explanation for nested structures.

Nestedness was measured using the spectral radius (*Methods*), an accurate and reliable approach that correlates strongly with other recommended metrics [6, 17]. As the capacity for nested structures is tied to the fill of a network (the fraction of realized links), we scaled nestedness values for networks produced by the phenology and degree distribution models between corresponding values for Erdös-Rényi random graphs and empirical networks ($\Delta\tilde{\rho}$; *Methods*).

Emergence of nestedness

At all levels of temporal aggregation (week, month, year), empirical networks were more nested than Erdös-Rényi random graphs, as were networks produced by the degree distribution and phenology models (Fig. 2). Interestingly, when daily interaction data were aggregated to weekly networks, phenology and degree distribution models generated similar levels of nestedness, around a quarter of that for empirical networks. For monthly aggregated networks, the proportion of generated nestedness rose to roughly half for the phenology model, but remained at around a quarter for the degree distribution model. Not all of the 22 plant-pollinator data sets were suitable for constructing weekly and monthly networks (Supplementary Table S1), but in all cases we could reliably analyze data aggregated over an entire year, i.e., a single field season — the usual temporal resolution of published plant-pollinator networks. At the annual level, across all 24 data sets, the degree distribution model generated around 30% of empirically observed nestedness,

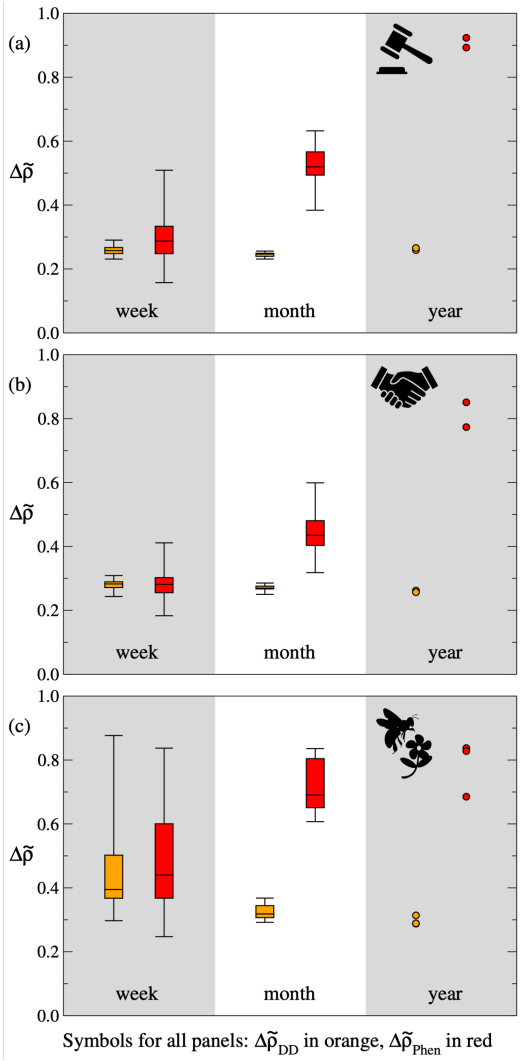

Figure 2: Nestedness generated by the degree distribution and phenology models applied to three data sets — buyer-seller transactions in auction (a) and negotiation (b) fish markets in Boulogne-sur-Mer, France, and plant-pollinator visitations in field sites at Rocky Mountain Biological Laboratory, Colorado, USA (c) — at three levels of temporal aggregation: week (left), month (middle), and year (right). Values for degree distribution model networks ($\Delta\tilde{\rho}_{\mathrm{DD}}$, orange) and phenology model networks ($\Delta\tilde{\rho}_{\mathrm{Phen}}$, red) are scaled between corresponding values for Erdös-Rényi random graphs ($\Delta\tilde{\rho} = 0$) and empirical networks ($\Delta\tilde{\rho} = 1$). Four outlier points have been omitted from the week box plots in panel (c): one for the degree distribution model ($\Delta\tilde{\rho}_{\mathrm{DD}} = 1.9$) and three for the phenology model ($\Delta\tilde{\rho}_{\mathrm{Phen}} = 1.1, 1.2, 5.5$). While the levels of nestedness generated by the degree distribution model remain relatively low as temporal aggregation increases, the phenology model generates values that approach empirically observed levels. Full time series for this figure can be found in Supplementary Fig. S1.

Table 1: Empirical network properties and proportion of empirical nestedness generated by the degree distribution ($\Delta\tilde{\rho}_{\mathrm{DD}}$) and phenology ($\Delta\tilde{\rho}_{\mathrm{Phen}}$) models at an annual level of temporal aggregation, mean and standard deviation (in parentheses) of 1000 realizations.

Data set	Ref	Rows	Cols	Fill	Days	Years	$\Delta\tilde{\rho}_{\mathrm{DD}}$	$\Delta\tilde{\rho}_{\mathrm{Phen}}$
Auction	[22]	195	100	0.38	525	2	0.26 (0.02)	0.91 (0.02)
Negotiation	[22]	207	93	0.50	538	2	0.26 (0.03)	0.81 (0.03)
CaraDonna	[27]	46	93	0.14	106	3	0.30 (0.09)	0.78 (0.10)
Alarcon2008	[28]	56	232	0.07	50	3	0.36 (0.08)	0.63 (0.08)
Benadi2014	[29]	121	406	0.04	46	1	0.27 (0.02)	0.83 (0.03)
Burkle2012	[30]	45	127	0.08	17	1	0.28 (0.05)	0.54 (0.05)
Chacoff2018	[31]	59	198	0.09	93	6	0.31 (0.09)	0.73 (0.10)
Fruend2010	[32]	103	160	0.06	51	1	0.29 (0.05)	1.03 (0.06)
Lara-Romero2016	[33]	22	146	0.16	22	1	0.30 (0.06)	0.82 (0.07)
LeBuhnYY	[23]	90	134	0.05	74	2	0.29 (0.07)	0.89 (0.08)
Olito2015	[34]	43	125	0.06	32	1	0.33 (0.08)	1.02 (0.09)
Rasmussen2013	[35]	39	108	0.10	106	2	0.31 (0.09)	1.01 (0.11)
ResascoYY	[23]	35	208	0.08	51	3	0.37 (0.10)	0.71 (0.11)
Simanonok2014	[36]	45	126	0.07	44	1	0.29 (0.06)	0.93 (0.07)
Thompson2018	[37]	41	127	0.09	37	1	0.35 (0.07)	0.99 (0.08)
Vazquez2003	[38]	14	117	0.13	76	1	0.44 (0.13)	0.92 (0.12)
Weiner2014	[39]	97	405	0.05	70	1	0.29 (0.02)	0.81 (0.02)
Winfree2014	[40]	59	85	0.09	33	2	0.31 (0.10)	0.67 (0.11)
WinfreeYYa	[41]	113	120	0.08	77	3	0.30 (0.08)	0.82 (0.09)
WinfreeYYb	[41]	112	143	0.07	81	3	0.29 (0.06)	0.83 (0.07)
WinfreeYYc	[23]	14	33	0.21	8	1	0.35 (0.14)	0.50 (0.14)
WinfreeYYd	[42]	101	95	0.12	54	1	0.26 (0.03)	0.80 (0.03)
WinfreeYYe	[23]	84	79	0.12	72	3	0.30 (0.07)	0.96 (0.10)
WinfreeYYf	[43]	57	84	0.08	22	1	0.27 (0.06)	0.54 (0.07)

compared to around 80% for the phenology model (Table 1). Thus, the proportion of nestedness explained by phenology increases as temporal data are aggregated, reaching close to empirically observed values as the level of aggregation matches a complete phenological cycle.

In order to see how closely model predictions matched empirical data, we used an assessment scheme that classified each link prediction into one of four categories: true positive, true negative, false positive, false negative (Fig. 1). From this information, we measured model accuracy using two summary statistics: the phi coefficient and the F-score (*Methods*). The phi coefficient is the equivalent of the Pearson correlation coefficient but for two binary variables rather than two continuous variables; for interaction networks, the binary variables are the presence or absence of each potential link in a model-generated network compared to its presence or absence in the empirical network. The phi coefficient provides a full-network assessment of model accuracy and is complemented by the F-score, which focuses on the subset of realized links (i.e., "positive" observations) in the empirical network. In the context of the phenology model, the F-score measures the relative contributions of actor co-presences (i.e., true positives) versus interaction preferences (departures from co-presence, i.e., false positives and false negatives) to network structure.

The phenology model outperformed the degree distribution model for all 24 data sets (Supplementary Table S2). Additionally, the amount of nestedness ($\Delta\tilde{\rho}$) generated by the phenology model was strongly correlated with the phi coefficient — Spearman's rank correlation coefficient, $r = 0.42$, $p = 0.04$, $n = 24$ — and to a lesser extent the F-score, $r = 0.29$, $p = 0.17$, $n = 24$. By contrast, for the degree distribution model, there was no correlation between nestedness and the phi coefficient, $r = -0.05$, $p = 0.82$, $n = 24$, nor nestedness and the F-score, $r = 0.03$, $p = 0.88$, $n = 24$. For the phenology model, at the annual level of temporal aggregation (the level at which all data sets could be compared), there was no notable difference in phi coefficients between fish market and plant-pollinator data sets. However, F-scores for the two fish market data sets, $F_1 = 0.67$ for the auction market and $F_1 = 0.68$ for the negotiation market, were much higher than the average value for the 22 plant-pollinator data sets, $F_1 = 0.33 \pm 0.08$ (mean $\pm$ standard deviation). For the annual plant-pollinator networks, this F-score result suggests that roughly one-third of links can be explained by the relative availability of interaction partners. The remaining two-thirds of links, which represent observed interactions between plants and pollinators that are rarely co-present (i.e., phenology model false negatives), may then be attributed either to opportunistic interactions or to preferences for particular interaction partners that are not routinely available.

Backbone of nested interactions

Interestingly, the links that were most readily explained by frequent actor co-presences appear to form a *backbone of nested interactions*. We investigated this possibility by comparing the relative levels of nestedness for the subset of true positive links (attributable to co-presences) to the subset of false negative links (attributable to interaction preferences). For 15 of the 24 data sets, we found that: (i) true positive links were significantly more nested than false negative links, and (ii) true positive links were significantly more nested relative to Erdös-Rényi random graphs than false negative links were to Erdös-Rényi random graphs (Supplementary Table S3). This finding indicates that the set of interactions generated by frequent actor co-presences is more nested than the other, remaining set of interactions, which are distributed more randomly among actors.

Discussion

Importance of time

In general, system dynamics play an integral role in structuring interaction networks. This influence ranges from straightforward and unambiguous, e.g., the impossibility of an interaction if two actors are never around at the same time, to complex and subtle, e.g., interaction partners co-evolving in ways that alter their ability to interact with other actors. Here, we explored the relationship between nestedness and how often actors overlap in space and time. We proposed the phenology model, which assumes the more often actors overlap the more likely they are to interact, and applied it to buyer-seller fish market networks and plant-pollinator visitation networks. In this way, we sought to understand nestedness in terms of potential generative mechanisms — temporal processes that drive the emergence of nestedness in empirical data — rather than topological properties that are known only *after* networks have been constructed.

Using daily interaction data, we found that the proportion of empirically observed nestedness generated by the phenology model increased with the level of temporal aggregation used to construct networks. By contrast, the performance of the degree distribution model remained relatively poor as the data aggregation window increased from week to month to year. A formal model selection procedure is difficult to specify due the different conceptual motivations for the two models and the different types of data required to parameterize them. For the degree distribution model, link probabilities are derived from the topology of a final time-aggregated network, whereas for the phenology model, link probabilities are based on the number of times actors are co-present within a given sampling period. For the data sets analyzed here, it is clear that the phenology model performed noticeably better than the degree distribution model at time scales longer than a week; nevertheless, in future studies it will still be worth considering both models (or variants of the two models) in parallel.

The closest agreement between the phenology model and empirical data was for annual networks, for which the model was also able to predict a large fraction of empirical links. There are both technical and scientific reasons for this result. For weekly networks, there are inherently little data for the phenology model to work with and therefore limited opportunities for it to produce network structures that differ significantly from Erdös-Rényi random graphs. At such short time scales, with the plant-pollinator data in particular, the phenology model is mostly picking up signals related to sampling effort and relative abundance, namely, that actor presences and absences in the model reflect high and low species abundances more so than phenological effects based on reliable estimates of co-presence. For monthly networks, the maximum number of possible co-presences increases from seven to 31 days, thereby providing more leeway to test the effect of co-presence frequency on network structure. For annual networks, data aggregation covers a complete phenological cycle (e.g., one financial year, one full flowering season) and offers the widest range of possible and recorded co-presence values.

In histograms of the relative frequency of co-presences at the three levels of temporal aggregation, while the median number of co-presences hardly increases, e.g., auction market: 1.1 ± 0.3 for weeks (mean $\pm$ standard deviation for 91 networks), 2.3 ± 0.5 for months (23 networks), and 8.5 ± 0.7 for years (2 networks), the right tail of the distribution becomes noticeably fatter, e.g., auction market largest co-presence values: 4.1 ± 1.0 for weeks, 13.2 ± 2.6 for months, and 117.5 ± 0.7 for years (Supplementary Figs. S2–S4). The increasing weight in the right tail of the co-presence distribution allows the phenology model to pick out a more consistent set of (nested) interactions. This analysis reiterates the importance of matching the level of temporal aggregation to the natural phenological time scale of a system to maximize the predictive power of simple temporal models like the phenology model.

Presence versus co-presence

The phenology model assumes the more often two actors are co-present the more likely they are to interact. It is reasonable to expect some actors are frequently co-present with other actors simply because they are generally more present during a given time period. Since it is easier to assess the relative presences of actors compared to collecting additional information on how often actors overlap through time, it is interesting to see how well presence, rather than co-presence, explains empirical nestedness. To do so, we designed the *presence model* by modifying the phenology model slightly: instead of scaling the probability of an interaction by the number of days two actors overlap, we scale the probability by the average number of days the actors were present during a specified time period (*Methods*).

The proportion of empirical nestedness generated by the presence model typically fell between values for the degree distribution and phenology models but was higher for the auction market data set and 6 of the 22 plant-pollinator data sets (Supplementary Table S4). At the annual level, across all 24 data sets, the presence model generated around 70% of empirically observed nestedness, compared to 30% and 80% for the degree distribution and phenology models, respectively. While the presence model generated relatively high nestedness for the auction market data set ($\Delta\tilde{\rho} = 1.04$), the value was much lower for the negotiation market data set ($\Delta\tilde{\rho} = 0.37$). However, link prediction accuracies were similar for the two fish market data sets, around 64% of the explanatory power of the phenology model for the phi coefficient and 88% for the F-score. This slightly worse performance extended to the plant-pollinator data sets, with the presence model offering around 80% of the explanatory power of the phenology model for the phi coefficient and 87% for the F-score (Supplementary Table S4).

Interestingly, while the proportion of empirical nestedness generated by the presence model differed notably between the two fish market data sets, prediction accuracies were similar. The lower nestedness ($\Delta\tilde{\rho}$) for the negotiation market data set suggests that predictions were concentrated on links between pairs of high-presence actors with similar numbers of interaction partners, at the expense of predictions involving high-presence actors with few interaction partners. On the other hand, the presence model generated higher nestedness for the auction market data set because, in addition to predicting interactions among a core of high-presence, high-degree actors, the model also predicted interactions between high-presence, low-degree actors and high-presence, high-degree actors (likely market makers). The fact that prediction accuracy was similar between auction and negotiation market data sets suggests that correct predictions were mainly restricted to links among a core set of high-presence, high-degree actors.

These findings extend our understanding of how co-occurrence patterns inform interaction patterns [44]. In particular, compared to the presence model, the phenology model improves link prediction accuracy on two fronts by: (i) refining which potential interactions among high-presence, high-degree actors are most likely to be realized based on how often they overlap through time and (ii) enabling more accurate predictions involving low-degree actors by accounting for how often they are co-present with high-degree actors, and not just how often they are present.

Future directions

Here, information on interaction partner availability through time was inferred from empirical data and implemented in the phenology model as simply the number of co-presences between each pair of actors. In future work, it will be worth exploring in detail how phenologies combine to produce nested structures. Accounting for exactly when and for how long actors are active is especially valuable for understanding the interactions of infrequent participants which, by definition, are associated with low co-presence values. These interactions are likely classified as false negatives, but considering phenologies more fully would help distinguish, for example, the genuine interactions of a rare plant with a long flowering period that is visited by a large number of short-lived pollinator species from the background noise of actor pairs that barely overlap.

The phenology model presents a viable path to a general mechanistic understanding of interaction network structure. Currently, the model explains the formation of backbones of nested interactions comprising around two-thirds of links in fish market networks and one-third of links in plant-pollinator networks; but this still leaves a large number of links that cannot be explained by actor co-presence frequency. Since the model assigns distinct probabilities to each potential link, systematically modifying these probabilities is one way of incorporating additional mechanisms. For plant-pollinator networks, this could include the effects of competition [45], phenotypic complementarity [10, 46], and interaction preferences [47]. For market-based networks, especially those in which buyers and sellers know one another and repeated transactions are common, loyalty is an prominent feature that impacts which interactions are likely to be realized [8, 22]. Interestingly, for the fish market data, auction networks were slightly more nested than negotiation networks. The phenology model also generated auction networks that were relatively more nested than negotiation networks, suggesting that co-presence had a greater impact on

nestedness when there was automatic matching of buyers and sellers through an order book (auction market) compared to when buyers and sellers made deals directly (negotiation market).

Since nestedness co-varies with other network metrics [48], it would be interesting to measure the extent to which an updated phenology model could explain the emergence of — and trade-offs between — other structural patterns, such as modularity and community structure [49, 50]. The fish market and plant-pollinator data sets both describe positive and mutually beneficial interactions between actors. It would be useful to contrast findings with data sets that include negative interactions, such as competition networks in which actors are negatively affected by the presence of other actors, or food webs in which one actor (a predator) benefits at the expense of another actor (a prey).

Although we have focused on binary networks, interaction partner availability and higher-level selection processes will also have an effect on the structure of quantitative or weighted networks that additionally document the relative strength or frequency of interactions [2]. Indeed, it may be even more important to account for actor timing when dealing with weighted interaction networks. For ecological systems, for example, phenology often dictates the abundance of species through regular growth and reproduction cycles, with the relative abundance of species, in turn, known to impact weighted network structure [6, 25]. Finally, it is worth noting that although the current study is focused on the temporal origin of nestedness in established and stable systems over relatively short and cyclical time frames, the phenology model would also be useful for studying interaction patterns as markets form and collapse and ecological communities assemble and disassemble [51–53].

Concluding remarks

For interaction networks, some influence of time — and of timing — on network structure is to be expected. A buyer and seller must agree on a price to make a trade, and a plant must be flowering for a pollinator to visit it. However, these short-term temporal dependencies are obscured in the time-aggregations commonly applied in traditional network analyses. In the case of plant-pollinator networks, this seemed acceptable since nestedness was thought to reflect constraints on trait matching shaped over evolutionary time scales. But we have now shown that phenology, which concerns much shorter, ecological time scales, can also lead to nested structures. It will therefore be important to revisit the relationship between nestedness and system-level properties such as feasibility, stability, robustness, and persistence, given the current uncertainty in how the two, temporally disconnected processes combine to generate nestedness and other structural features. For ecological interactions, anthropogenic drivers like climate change and urbanization are making phenological mismatches more likely, meaning there is an urgent need to trace the effects of shifting phenologies on network structure and the resulting impacts on species fitness and survival.

Materials and methods

Data. Auction and negotiation market data sets were obtained from an earlier study [22]. Data on daily buyer-seller transactions were collected over 21 months spanning two consecutive calendar years; we divided it into two parts comprising 10.5 months each, representing two "years" of daily buyer-seller transaction data. The 22 plant-pollinator data sets were previously compiled as part of an earlier study [23] and were selected for analysis here because they include visitations between individual insect pollinators and plants resolved to the daily level; source papers are listed in Table 1. Note that plant-pollinator data were not collected for all days in the overall sampling period of each data set; similarly, buyer-seller data excluded days when the market was closed, e.g., Sundays and national holidays. Temporally aggregated networks were built from consecutive, non-overlapping blocks of daily interaction data and only included nodes and links corresponding to buyer and seller individuals or plant and pollinator species that interacted during a given block of time. Multiple interactions in a single block were counted only once, leading to binary, interaction presence-absence networks.

Models. We considered four probabilistic models that conserve, on average, the total number of links in a time-aggregated empirical network. The Erdös-Rényi random graph and degree distribution model are well-established, while the phenology and presence models are newly presented here. Each model describes links as independent Bernoulli random variables with probabilities defined below.

Erdös-Rényi random graph. Each link has the same probability of being realized, $p_{ij}^{\mathrm{ER}} = L/(RC)$ (the fill or connectance of an empirical network); where L is the number of links in an empirical network, R is the number of nodes in the actor category listed as rows, e.g., sellers or plants, C is the number of nodes in the actor category listed as columns, e.g., buyers or pollinators, and i and j index nodes in the row category and column category, respectively.

Degree distribution model. Link probability is $p_{ij}^{\mathrm{DD}} = k_i/(2C) + k_j/(2R)$; where k_i is the empirical degree of node i in the row category and k_j is the empirical degree of node j in the column category. In addition to the total number of links in an empirical network, the degree distribution model also conserves the expected number of links associated with each node and therefore the expected degree distribution of the empirical network [18].

Phenology model. The model requires two parameters that are not derived from the topology of an empirical network. First is the sampling period, f, which is the number of days worth of interaction data included in the time-aggregated empirical network (if a different minimal temporal unit for resolving interaction data is used then f is specified in terms of this unit, e.g., hours instead of days). The second parameter is the co-presence matrix, N, that lists the number of instances — in terms of days or some other minimal temporal unit — each pair of actors overlap in space and time during the time-aggregated sampling period. If the probability of an interaction given a co-presence is p and the probability of no interaction is $q = 1 - p$, then the probability of at least one interaction during the sampling period and therefore the probability of a link is $p_{ij}^{\mathrm{Phen}} = 1 - q^{n_{ij}}$; where n_{ij} is the number of co-presences between actors represented by nodes i and j. In this way, the greater the number of co-presences, the higher the probability of a link.

Note that we assume a single probability of an interaction given a co-presence (i.e., p), which we set to a value such that the expected number of links equals the number of links in the empirical network. Since there are at most $f + 1$ distinct entries in N, ranging from 0 (i.e., actors never overlap during the sampling period) to f co-presences (i.e., actors always overlap), we find $p = 1 - q$ by solving the polynomial, $\sum_{k=0}^{k=f} x_k q^k - Z = 0$; where x_0 is the number of 0s in N, x_1 the number of 1s, etc., and $Z = RC - L$ is the number of zeros in the incidence matrix representation of the empirical network.

For the example in Fig. 1, $f = 6$ days and the co-presence matrix, N, is shown in panel b. Solving the corresponding polynomial, $q^0 + 2q + 4q^2 + 2q^3 - 3 = 0$, gives $q = 0.47$ and $p = 1 - 0.47 = 0.53$. Individual link probabilities in panel c can then be calculated from these values of p and q.

Presence model. Based on the phenology model, the presence model is designed to provide a comparison of the effect of actor presence on network structure relative to the effect of actor co-presence. Instead of the number of days each pair of actors overlap, the average number of days each pair of actors is present during the time-aggregated sampling period is listed in matrix N (in practice, the average number of days is rounded up to the nearest integer to ensure the polynomial described above is solvable). Everything else remains the same as in the phenology model. Note, however, that the interpretation of p as the probability of an interaction given a co-presence no longer holds, and therefore the presence model should be seen as a way to scale link probabilities according to actor prevalence or commonness, i.e., presence, rather than opportunities to interact, i.e., co-presence.

Measuring nestedness. A network can be represented as an adjacency matrix in which the full set of nodes is listed as both rows and columns and non-zero entries indicate the presence of a link between two nodes. For bipartite networks, there will be two blocks of zeros on the diagonal since there are no links between nodes in the same category; the (symmetric) off-diagonal block contains all relevant topological information and is known as the incidence matrix, as in the upper box of Fig. 1. Measuring nestedness involves computing the spectral radius — the absolute value of the dominant or largest eigenvalue — of the adjacency matrix [6, 17, 54]. For a given size and fill, i.e., total number of nodes and links, the larger the spectral radius the more nested the particular network topology. To enable comparison between different networks, raw values of the spectral radius can be normalized by their maximum possible values [55] to give a measure for nestedness that ranges between ~ 0 and 1: $\rho = \rho_{\mathrm{raw}}/\sqrt{2L - R - C + 1}$.

We used normalized spectral radius values of Erdös-Rényi random graphs as a baseline for contextualizing the relative nestedness of empirical networks and more complex model-generated networks (here, the degree distribution and phenology models): $\Delta\rho = \rho - \rho_{\mathrm{ER}}$; where ρ_{ER} is the mean value of the normalized spectral radius of an ensemble of Erdös-Rényi random graphs (here, we used 10,000 networks) parameterized according to the the number of nodes and links in the focal empirical network. Empirical or model-generated networks with $\Delta\rho > 0$ are nested, while networks with $\Delta\rho < 0$ are anti-nested [6]. To more easily compare models and their ability to generate empirically observed levels of nestedness, we considered a further rescaling: $\Delta\tilde{\rho} = (\rho - \rho_{\mathrm{ER}})/(\rho_{\mathrm{emp}} - \rho_{\mathrm{ER}})$; where ρ_{emp} is the normalized spectral radius of the focal empirical network. A model-generated network with $\Delta\tilde{\rho} = 0$ has the same level of nestedness as a corresponding Erdös-Rényi random graph, while $\Delta\tilde{\rho} = 1$ indicates that the model-generated network has the same level of nestedness as the empirical network.

Assessing link prediction accuracy. We used the confusion matrix, a two dimensional contingency

table, to assess the accuracy of model predictions [56]. For a given model-generated network, each entry in its corresponding incidence matrix can be assigned to one of four categories: true positive (TP) — link present in both model and empirical networks; true negative (TN) — link absent in both model and empirical networks; false positive (FP) — link present in model network but absent in empirical network; false negative (FN) — link absent in model network but present in empirical network. We summarized model accuracy using two well-established statistics: the phi coefficient and the F-score [57].

phi coefficient. This statistic involves all four categories and is the equivalent of the Pearson correlation coefficient for two binary variables: $\phi = \frac{\mathrm{TP}\times\mathrm{TN}-\mathrm{FP}\times\mathrm{FN}}{\sqrt{(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}}$. Values for the phi coefficient range from -1 (perfect disagreement) through 0 (no relationship) to 1 (perfect agreement). Note that since we constrain model-generated networks to have, on average, the same number of links as in an empirical network, $\mathrm{FP} \approx \mathrm{FN}$ and therefore phi coefficient values in this case approximate values for the informedness or Youden's J statistic, $J = \text{sensitivity} + \text{specificity} - 1 = \mathrm{TP}/(\mathrm{TP}+\mathrm{FN}) + \mathrm{TN}/(\mathrm{TN}+\mathrm{FP}) - 1$.

F-score. This statistic focuses on the subset of realized links in an empirical network and as such does not involve true negatives: $F_1 = 2\mathrm{TP}/(2\mathrm{TP}+\mathrm{FP}+\mathrm{FN})$. Values for the F-score range from 0 (no links correctly predicted) to 1 (all links correctly predicted). Since $\mathrm{FP} \approx \mathrm{FN}$, F-score values in this case approximate values for the sensitivity or true positive rate, $\mathrm{TPR} = \mathrm{TP}/(\mathrm{TP}+\mathrm{FN})$. It is worth noting that, for Erdös-Rényi random graphs, sensitivity values (and therefore F-score values, here) approximate the fill or connectance, L/RC, of an empirical network.

Backbone of nested interactions. We tested the hypothesis that the subset of true positive links predicted by the phenology model are more nested than the subset of false negative interactions in two parts. First, we calculated the ratio of nestedness values for the subsets of true positive and false negative links; if $\rho_{\mathrm{TP}}/\rho_{\mathrm{FN}} > 1$, then interactions between frequently co-present actors are more nested than the remaining interactions. To account for any confounding size-based effects, we also compared the deviations of nestedness from comparable Erdös-Rényi random graphs. Specifically, we calculated the proportion of excess nestedness for the phenology model relative to the maximum possible value: $\rho_{\mathrm{TP},\Delta\mathrm{ER}} = (\rho_{\mathrm{TP}} - \rho_{\mathrm{TP,ER}})/(1-\rho_{\mathrm{TP,\ ER}})$ for the subset of true positive links and $\rho_{\mathrm{FN,ER}} = (\rho_{\mathrm{FN}} - \rho_{\mathrm{FN,ER}})/(1-\rho_{\mathrm{FN,ER}})$ for the subset of false negative links. If $\rho_{\mathrm{TP},\Delta\mathrm{ER}}/\rho_{\mathrm{FN},\Delta\mathrm{ER}} > 1$, then interactions between frequently co-present actors are more nested relative to Erdös-Rényi random graphs than the remaining interactions.

Data availability. All data analyzed in this study have been published previously and are available through the references in Table 1. Code to perform analyses are available on GitHub in the public repository, https://github.com/pstaniczenko/TemporalOriginNestedness.

Author contributions. P.P.A.S. and D.P. designed research, performed research, analyzed data, and wrote the paper; P.P.A.S. also contributed new analytical tools.

Funding statement. There were no funders for this work.

[1] B. Patterson, W. Atmar, *Biological Journal of the Linnean Society* **28**, 65 (1986).

[2] M. Mariani, Z.-M. Ren, J. Bascompte, C. Tessone, *Physics Reports* **813**, 1 (2019).

[3] S. Bustos, C. Gomez, R. Hausmann, C. Hidalgo, *PLOS ONE* **7**, e49393 (2012).

[4] L. Bascompte, P. Jordano, C. Melián, J. Olesen, *Proceedings of the National Academy of Sciences USA* **100**, 9383 (2003).

[5] J. Grilli, *et al.*, *Nature Communications* **8**, 14389 (2017).

[6] P. Staniczenko, J. Kopp, S. Allesina, *Nature Communications* **4**, 1931 (2013).

[7] R. Rohr, S. Saavedra, J. Bascompte, *Science* **345**, 1253497 (2014).

[8] R. Kranton, D. Minehart, *American Economic Review* **91**, 485 (2001).

[9] U. Bastolla, *et al.*, *Nature* **458**, 1018 (2009).

[10] E. Rezende, P. Jordano, J. Bascompte, *Oikos* **116**, 1919 (2007).

[11] K. Takemoto, M. Arita, *Journal of Theoretical Biology* **264**, 782 (2010).

[12] S. Valverde, *et al.*, *Nature Ecology and Evolution* **2**, 94 (2018).

[13] P. Milgrom, *Putting auction theory to work* (Cambridge University Press, 2004).

[14] L. Santamaría, M. Rodríguez-Gironés, *PLOS Biology* **5**, e31 (2007).

[15] S. Saavedra, F. Reed-Tsochas, B. Uzzi, *Nature* **457**, 463 (2009).

[16] S. Suweis, F. Simini, J. Banavar, A. Maritan, *Nature* **500**, 449 (2013).

[17] C. Payrató-Borràs, L. Hernández, Y. Moreno, *Ecology and Evolution* **10**, 11906 (2020).

[18] W. Ulrich, M. Almeida-Neto, N. Gotelli, *Oikos* **118**, 3 (2009).

[19] W. Ulrich, N. Gotelli, *Ecology* **88**, 1824 (2007).

[20] L. Joppa, J. Montoya, R. Solé, J. Sanderson, S. Pimm, *Evolutionary Ecology Research* **12**, 35 (2010).

[21] C. Payrató-Borràs, L. Hernández, Y. Moreno, *Physical Review X* **9**, 031024 (2019).

[22] L. Hernández, A. Vignes, S. Saba, *PLOS ONE* **13**, e0196206 (2018).

[23] B. Schwarz, *et al.*, *Oikos* **129**, 1289 (2020).

[24] F. Encinas-Viso, T. Revilla, R. Etienne, *Ecology Letters* **15**, 198 (2012).

[25] P. CaraDonna, *et al.*, *Ecology Letters* **20**, 385 (2017).

[26] B. Mora, E. Shin, P. CaraDonna, D. Stouffer, *Nature Communications* **11**, 4086 (2020).

[27] P. CaraDonna, *Environmental Data Initiative, https://doi.org/10.6073/pasta/27dc02fe1655e3896f20326fed5cb95f* (2020).

[28] R. Alarcón, N. Waser, J. Ollerton, *Oikos* **117**, 1796 (2008).

[29] G. Benadi, T. Hovestadt, H.-J. Poethke, N. Blthgen, *Journal of Animal Ecology* **83**, 639 (2014).

[30] L. Burkle, T. Knight, *Ecology* **93**, 2329 (2012).

[31] N. Chacoff, J. Resasco, D. Vázquez, *Ecology* **99**, 21 (2018).

[32] J. Fründ, C. Dormann, T. Tscharntke, *Ecology Letters* **14**, 896 (2011).

[33] C. Lara-Romero, C. García, J. Morente-López, J. Iriondo, *Functional Ecology* **30**, 1521 (2016).

[34] C. Olito, J. Fox, *Oikos* **124**, 428 (2015).

[35] C. Rasmussen, Y. Dupont, J. Mosbacher, K. Trøjelsgaard, J. Olesen, *PLOS ONE* **8**, e81694 (2013).

[36] M. P. Simanonok, L. A. Burkle, *Ecosphere* **5**, 149 (2014).

[37] A. Thompson, T. Knight, *Oecologia* **187**, 135 (2018).

[38] D. Vázquez, D. Simberloff, *Ecology Letters* **6**, 1077 (2003).

[39] C. Weiner, M. Werner, K. Linsenmair, N. Blüthgen, *Ecology* **95**, 466 (2014).

[40] R. Winfree, N. Williams, J. Dushoff, C. Kremen, *The American Naturalist* **183**, 600 (2014).

[41] M. MacLeod, *et al.*, *Journal of Applied Ecology* **57**, 413 (2020).

[42] M. Roswell, J. Dushoff, R. Winfree, *PLOS ONE* **14**, e0214909 (2019).

[43] C. Smith, L. Weinman, J. Gibbs, R. Winfree, *Journal of Animal Ecology* **88**, 1158 (2019).

[44] A. Barner, K. Coblentz, S. Hacker, B. Menge, *Ecology* **99**, 557 (2018).

[45] F. Valdovinos, *et al.*, *Ecology Letters* **19**, 1277 (2016).

[46] E.-L. Marjakangas, *et al.*, *Ecological Monographs* **92**, e1502 (2021).

[47] P. Staniczenko, *et al.*, *Nature Communications* **8**, 792 (2017).

[48] S. Jonhson, V. Domínguez-García, , M. Muñoz, *PLOS ONE* **8**, e74025 (2013).

[49] M. Fortuna, *et al.*, *Journal of Animal Ecology* **79**, 811 (2010).

[50] S. Fortunato, D. Hric, *Physics Reports* **659**, 1 (2016).

[51] S. Saavedra, R. Rohr, J. Olesen, J. Bascompte, *Ecology and Evolution* **6**, 997 (2016).

[52] L. Ponisio, *et al.*, *Frontiers in Ecology and Evolution* **7** (2019).

[53] L. Becker, N. Blüthgen, B. Drossel, *American Naturalist* **200**, 303 (2022).

[54] F. Bell, D. Cvetković, P. Rowlinson, S. Simić, *Linear Algebra and its Applications* **429**, 2168 (2008).

[55] H. Yuan, *Linear Algebra and its Applications* **108**, 135 (1988).

[56] T. Fawcett, *Pattern Recognition Letters* **27**, 861 (2006).

[57] D. Powers, *Journal of Machine Learning Technologies* **2**, 37 (2011).