

Article

Constructing ecological interaction networks by correlation analysis: hints from community sampling

WenJun Zhang

School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China; International Academy of Ecology and Environmental Sciences, Hong Kong

E-mail: zhwj@mail.sysu.edu.cn, wjzhang@iaees.org

Received 6 January 2011; Accepted 12 March 2011; Published online 1 September 2011

IAEES

Abstract

A set of methodology for constructing ecological interaction networks by correlation analysis of community sampling data was presented in this study. Nearly 30 data sets at different levels of taxa for different sampling seasons and locations were used to construct networks and find network properties.

I defined the network constructed by Pearson linear correlation is the linear network, and the network constructed by quasi-linear correlation measure (e.g., Spearman correlation) is the quasi-linear network. Two taxa with statistically significant linear or quasi-linear correlation are determined to interact. The quasi-linear network is more general than linear network.

The results revealed that correlation distributions of Pearson linear correlation and partial linear correlation constructed networks are unimodal functions and most of them are short-head (mostly negative correlations) and long-tailed (mostly positive correlations). Spearman correlation distributions are either long-head and short-tailed unimodal functions or monotonically increasing functions. It was found that both mean partial linear correlation and mean Pearson linear correlation were approximately 0. The proportion of positive (partial) linear correlations declined significantly with the increase in taxa. The mean (partial) linear correlation declined significantly with the increase of taxa. More than 90% of network interactions are positive interactions. The average connectance was 9.8% (9.3%) for (partial) linear correlation constructed network. The parameter λ in power law distribution ($L(x)=x^{-\lambda}$) increased as the decline of taxon level (from functional group to species) for the partial linear correlation constructed network. λ is in average 0.8 to 0.9. The number of (positive) interactions increased with the number of taxa for both linear and partial linear correlations constructed networks. The addition of a taxon would result in an increase of 0.4 (0.3) interactions (positive interactions) in the partial linear correlation constructed network. And the addition of a taxon would result in an increase of 3 interactions (positive interactions) in the linear correlation constructed network. For partial linear correlation constructed network, the network connectance decreased as the number of taxa. The constant connectance hypothesis did not hold for our networks. It was found that network structure changed with season and location. The same taxon in the network would connect to different taxa as the change of season and location. A higher level of species aggregation may be used to find a more stable network structure.

Positive interactions were considered to be caused mainly by mutualism, predation/parasitism, etc. the number and portion of positive interactions may be the most important indices for community stability and functionality. Mutualism is the most significant trophic relationship, seconded by predation/parasitism, and competition is the worst for community stability.

Keywords linear network; quasi-linear network; network construction; community sampling; correlation analysis; degree distribution; link; interaction.

1 Introduction

So far a lot of studies have been done for simple networks. Many results and methods were obtained from those studies. However, the networks met in last decade become more and more complex. There are always large numbers of vertices and links in a complex network. It will be impossible to approach such networks by using classical methods or algorithms. Graph theory, combinatorial optimization, statistics, and stochastic processes, etc., are thus becoming the scientific basis and effective tools for studying complex networks (Ferrarini, 2011; Zhang, 2011a; Zhang, 2012). It has been found that in the random network, degree distribution is binomial distribution, and its limit model is Poisson distribution. In a random network, the majority of vertices have the same degree with the average. In the complex network, degree distribution is a power law distribution, and the network is called a scale-free network (Barabasi and Albert, 1999; Barabasi, 2009). Many complex networks, such as Internet, metabolic networks, communication networks, etc., are scale-free networks. A property of the scale-free networks is that the structure and the evolution of networks are inseparable. Scale-free networks constantly change because of the arrival of nodes and links (Barabasi and Albert, 1999).

A food web usually contains a large number of species, in which they interact with each other by direct or indirect interactions (Schoener, 1993; May, 1983). Interactions always occur at different levels of taxa and spatial scales (Schmitz and Booth, 1997).

The definition and classification of entities (taxa) and environment are important. We may sample some designated area and document all such interactions to construct a network within the area and the sites sampled (Butts, 2009; Schoenly and Zhang, 1999a,b). The approaches used will completely affect the resultant network structure, and the robustness of the network. Simultaneous analysis of the same system at multiple levels of aggregation is suggested (Butts, 2009).

Most studies on food webs so far are based on species. Species aggregation, such as family and functional group, are seldom treated in food web studies. In a sense, hierarchical species aggregation would better represent biodiversity and environmental stability. If the environment for a higher hierarchical taxon are not suitable for survive this taxon, then its members at the lower hierarchical taxon may not survive in this environment. Ecological networks based on diverse hierarchies of taxa are thus necessary. There are more than two species at the higher hierarchical taxon and various species in the taxon may play different roles. Some of them are predators and others may be preys. So it is hard to determine the orientation of a link, which will result in an un-oriented graph (network). For these networks, correlation analysis above is an effective method to determine between-taxon interactions.

It was found that scale and resolution affect food web structure (Martinez, 1993a,b). Co-extinction tends to involve taxonomically related species (phylogenetic species, e.g. species belonging to the same genera) (Rezende et al., 2007). Trophic taxa extinct more rapidly (Petchey et al., 2008). Phylogenetic relatedness can partly explain the patterns of interactions between species (Rezende et al., 2007).

Food webs are found to be robust to the random extinction of species but rely on a few well-connected species that act as glue keeping the whole network together. If these key species disappear, it is expected that the entire network will collapse very rapidly (Memmott et al., 2004; Montoya et al., 2006; Dunne et al., 2002).

A key challenge in quantifying interspecific variation within diverse plant communities is that many species occur at extremely low densities, making it infeasible to collect sufficient data for meaningful statistical analyses at the species level for most species (Comita et al., 2010). Previous studies of density dependence have dealt with this by limiting analyses to the most abundant species in the community or lumping species into broad functional groups or abundance classes. However, different patterns of species

aggregation that generate taxa can produce networks of different structural features. They influence the number of links per taxon, average length of food chain, etc. (Sugihara et al., 1997; Solow and Beet, 1998). To avoid misleading conclusions, the set of nodes should be defined so as to include all distinct entities that are capable of participating in the relationship under study (Butts, 2009). Although there are some shortcomings with species aggregation, however these approaches can reduce various deviations (Williams and Martinez, 2000) and are thus widely accepted in the network studies (Arii et al., 2007). So far potential influences of species aggregation on degree distribution of food web have not been approached (Arii et al., 2007).

Past ecological network studies always omitted the interaction strength (Paine, 1980, 1988, 1992). However, many forms of interaction are inherently episodic and occur at variable rates (Whitehead and Dufault, 1999). Just imagine two arthropod networks with the same species and known links, theoretically their food webs are the same. However we find that there are huge differences between two networks, such as distribution of interaction strength, community functionality, etc. Thus static food webs are not able to represent actual ecological networks. Dichotomization of such data not only obscures such variation but also requires selecting a threshold level, the choice of which can substantially alter the properties of the resulting network, both directly through selective tie removal (Onnela et al., 2007) and indirectly through changes in network density (Faust, 2007).

Network structures with different connection strengths can vary greatly. It is necessary to determine whether the relationship under study is sufficiently stable over the period of interest (Butts, 2006, 2009). For relations known to be highly heterogeneous, no single threshold may suffice and a weighted graph representation will frequently be more appropriate. To assess the effectiveness of such approximations and provide concrete, empirically validated guidelines for practice within particular problem domains would be a welcome addition to the literature (Butts, 2006, 2009). The identification of common architectures, robust in the face of perturbations regardless of specific details may emerge from such studies (Matsuda and Namba, 1991; Bascompte, 2009).

Methodology for constructing ecological networks by correlation analysis of community sampling data were presented in this study. Nearly 30 data sets at different levels of taxa (species, family, functional group, etc.) for different sampling years and seasons and different countries were used to construct ecological networks and find network properties.

2 Materials and Methods

2.1 Sampling data

Totally 60 plots, each with 1 m² of rice field, were randomly sampled for arthropods using a machine sucker in Guangzhou, China, at September 16 and 30, and October 14 and 28 of 2006, respectively. Arthropods were taken to laboratory and identified to families and functional groups (herbivores, neutral arthropods, predators and parasitoids/parasites). Data were stored as 4 sample-by-family matrices and 4 sample-by-functional groups matrices.

In total of 80 plots, each with 1 m² of rice field, were sampled for arthropods using a machine sucker in Guangzhou, Zhuhai, Zhongshan and Dongguan cities of China at earlier September 2008. Arthropods were taken to laboratory and identified to species and families. Data were stored as a sample-by-species matrix and a sample-by-family matrix.

In Luzon of Philippines, arthropod samples for invertebrates were collected in the rice field on March, April, September and October, mid-1990s, respectively. In total of 60 samples were collected for each of four sampling dates. Invertebrates were sorted to stage (immatures, adults) and then identified to lowest possible

taxon. Data from the records were stored as 4 sample-by-species matrices (immatures and adults were listed separately, and defined as different (trophic) species in present study), and then lumped into 4 sample-by-family matrices, 4 functional group (in different seasons there were 20-21 functional groups respectively) matrices, and 4 macro-functional group (there are 7 macro-functional groups) matrices using LUMP method (Schoenly and Zhang, 1999a; Zhang, 2007a).

Following Schoenly and Zhang (1999), for data recorded in Luzon, a functional group was assigned to each taxon (adult and immature) based on what it was doing the majority of the time in the rice ecosystem, in which there were 7 macro functional groups and 35 functional groups:

(1) Herbivores: pollen feeder; external plant feeder; leaf roller/webber; case bearer and caseworm; leaf miner; gall former; borer; root feeder; shredder (aquatic); mixed (combination of two or more of above);

(2) Predators: terrestrial flyer; terrestrial crawler, walker, jumper, or hunter; neustonic (water surface) swimmer (semiaquatic); planktonic (water column) swimmer and diver; terrestrial web-builder; mixed (combination of two or more of above);

(3) Parasitoids/parasites: idiobiont (neuropteran ectoparasitoid); idiobiont (hymenopteran ectoparasitoid); koinobiont (hymenopteran endoparasitoid); koinobiont (dipteran endoparasitoid); parasite (aquatic or terrestrial); terrestrial blood sucker; koinobiont (strepsipteran endoparasitoid); flying adult that is searching, ovipositing, or larvipositing; idiobiont (acarine ectoparasitoid); koinobiont (nematode endoparasitoid); idiobiont (dipteran ectoparasitoid);

(4) Detritivores: collector (filterer, suspension feeder); collector (gatherer, deposit feeder); shredder, chewer of coarse particulate Matter;

(5) Tourists: tourist (nonpredatory species with no known functional role other than as prey in ecosystem);

(6) Omnivores: herbivore and predator; herbivore and detritivore; herbivore, predator, and detritivore;

(7) Dual insectivores: predator and parasitoid.

All data sets were tested for sample homogeneity (Zhang, 2011b) and used for further analysis.

2.2 Correlation analysis

2.2.1 Pearson linear correlation

2.2.2 Partial linear correlation

Partial linear correlation (R_{ij}) is based on Pearson linear correlation, which reflects between-taxon direct (pure, or net) correlation (Zhang, 2007b; Zhang, 2012).

2.2.3 Spearman correlation

See Schoenly and Zhang (1999b), and Zhang (2012) for Spearman correlation. Follow the calculation of partial linear correlation, I used Spearman partial correlation in the same way.

Between-taxon correlation can be used to find a network. If the taxon is not a species but family or functional group, etc, the relationship between two taxa will be complex because between-species interactions may be diverse. In this case the correlation measure is an alternative method to detect a network.

I define that the networks derived from linear correlation are linear networks, and the networks derived from Spearman correlation, which is a quasi-linear correlation measure, are a kind of quasi-linear networks. Statistically significant correlations, i.e., taxa pairs with significant linear or quasi-linear dependency are included in two kinds of networks respectively.

In a linear network, the states of two linked taxa will show a linear dependent relationship whereas for a quasi-linear network (created by partial Spearman correlation), the states of two linked taxa will show a quasi-linear dependent relationship. Taxa that never follow linear and quasi-linear relationships are excluded from the two networks respectively. In a sense, taxa in the network are easily predictable and taxa excluded from the network are hard to be predicted.

Using partial correlation will result in a network with all links as direct interactions. Pearson linear correlation and Spearman correlation can be used to create a network with links as indirect interactions. In later case an interaction is indirect.

Here I define an index, network compactness, to measure the interaction intensity in a network: network compactness =the sum of between-taxon correlation coefficients of interactions/the number of taxa. Two networks or ecosystems with the same type can be compared for their stability and maturity.

2.3 Degree and links

In present study, the number of links of a taxon is the degree of the taxon, and degree=(incoming-degree)+(outgoing-degree). Connectance of a network represents the connection intensity of the network. Connectance=actual interactions/potential interactions.

We avoid the assumption of dichotomous relationships by allowing between-taxon links (edges) to carry different weights (connection strength) (Butts, 2009). Abstractly all taxa in a system link each other with different link weights. The links with zero weight are deleted. Moreover, links with weights lower than desired can also be deleted. So the threshold of link weights may be defined to finally create a network.

Between-taxon correlation is a kind of link weight. A between-taxon correlation was treated as a link, or a directed or indirected interaction if the correlation is statistically significant. If the environment conditions for sampling arthropods are the same, spatial sampling may be used to substitute for temporal sampling and dynamic interactions can be represented by spatial changes of interactions. The statistically significant correlations should thus represent the true interactions (directed interactions for partial linear correlations) and the corresponding taxa pairs may be considered to have interactions or links. Moreover, the general principles of ecological interaction networks can be drawn by correlation analysis of community sampling data across various seasons and locations and taxon hierarchies.

3 Distribution of Between-taxon Correlations

Distributions of between-taxon correlation coefficients were calculated based on three correlation measures, Pearson linear correlation, partial linear correlation and Spearman correlation (Fig.1).

Fig.1 demonstrates that all of Pearson linear correlation and partial linear correlation based distributions are unimodal functions. Most of them are short-head (mostly negative correlations) and long-tailed (mostly positive correlations) unimodal curves (Fig.1). However, Spearman correlation based distributions show different patterns. They are either long-head and short-tailed unimodal functions or monotonically increasing functions (Fig.1). Most between-taxon Spearman correlations are positive correlations.

To describe Spearman correlation and Pearson linear (partial linear) correlation based distributions, I present here a distribution function as the following:

$$\begin{aligned} f(r) &= a(r-a)^{\beta} (b-r)^{\eta}, \quad a \leq r \leq b; \\ f(r) &= 0, \quad r > a \text{ or } r < b; \\ a &> 0, \beta \geq 0, \eta \geq 0. \end{aligned}$$

where $f(r)$: probability density (or frequency) function; a : scale parameter; β, η : shape parameters; a, b : position parameters. The probability density function is illustrated in Fig.3. Theoretically it may soundly describe all distributions indicated in Fig.3.

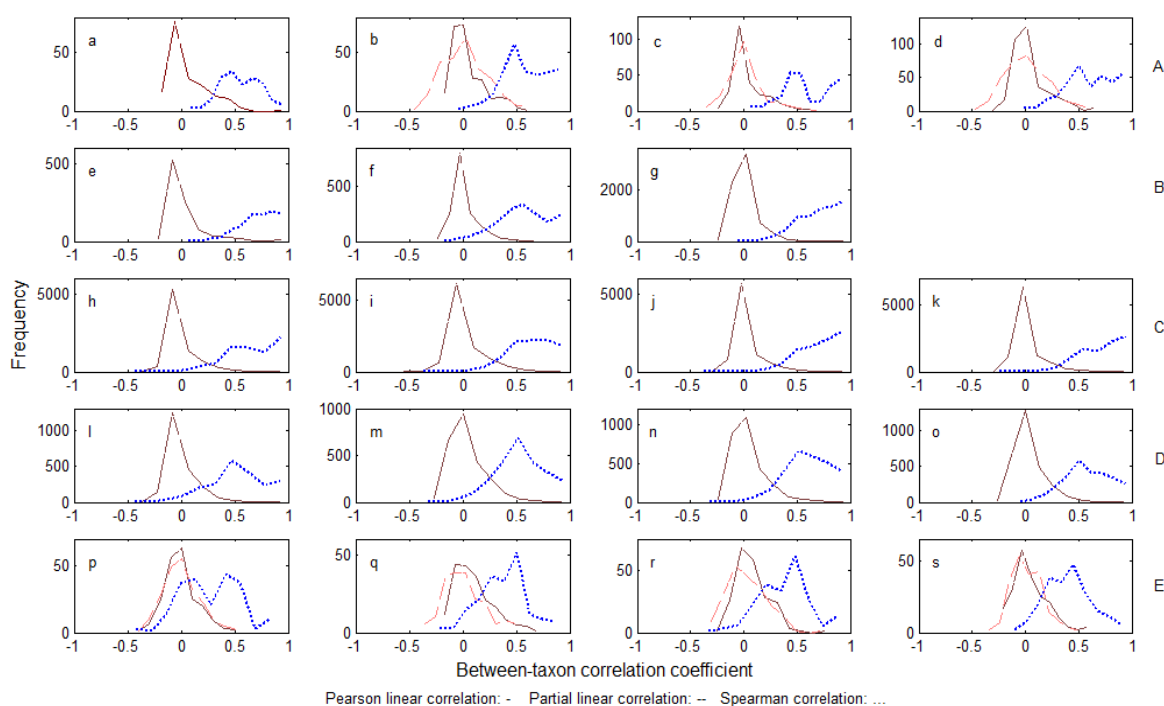
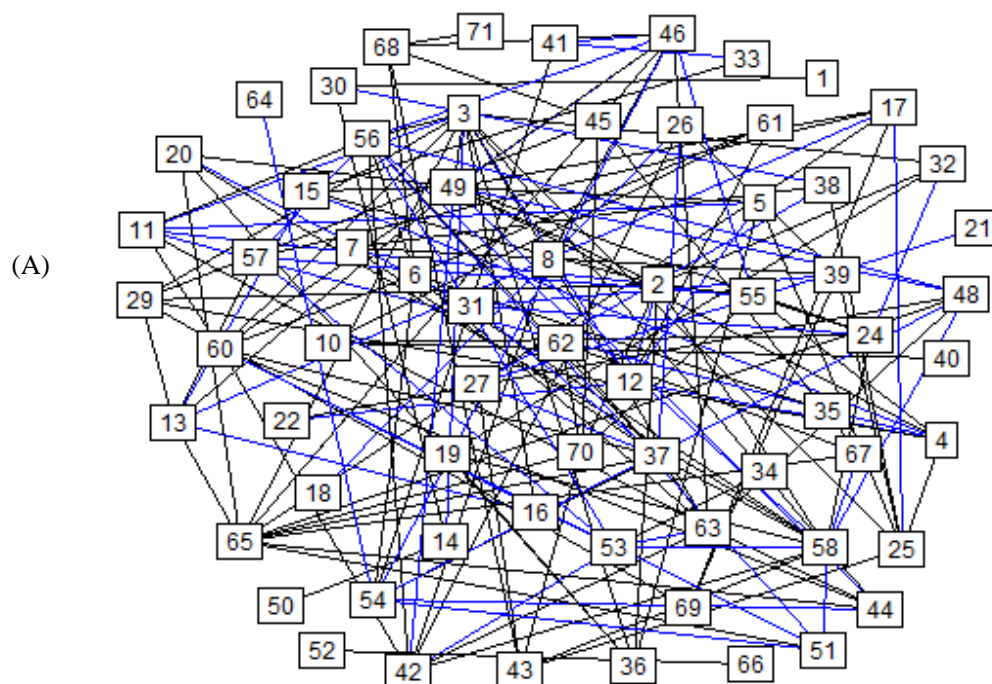


Fig. 1 Distributions of between-taxon correlation coefficients. A: distributions of between-family correlation coefficients for four seasons, China, 2006; B: distributions of between-family correlation coefficients for April and September and between-species correlation coefficients for September, China, 2008; C, D and E: distributions of between-species, between-family and between-functional group correlation coefficients for four seasons, Philippines (Fig. 2).



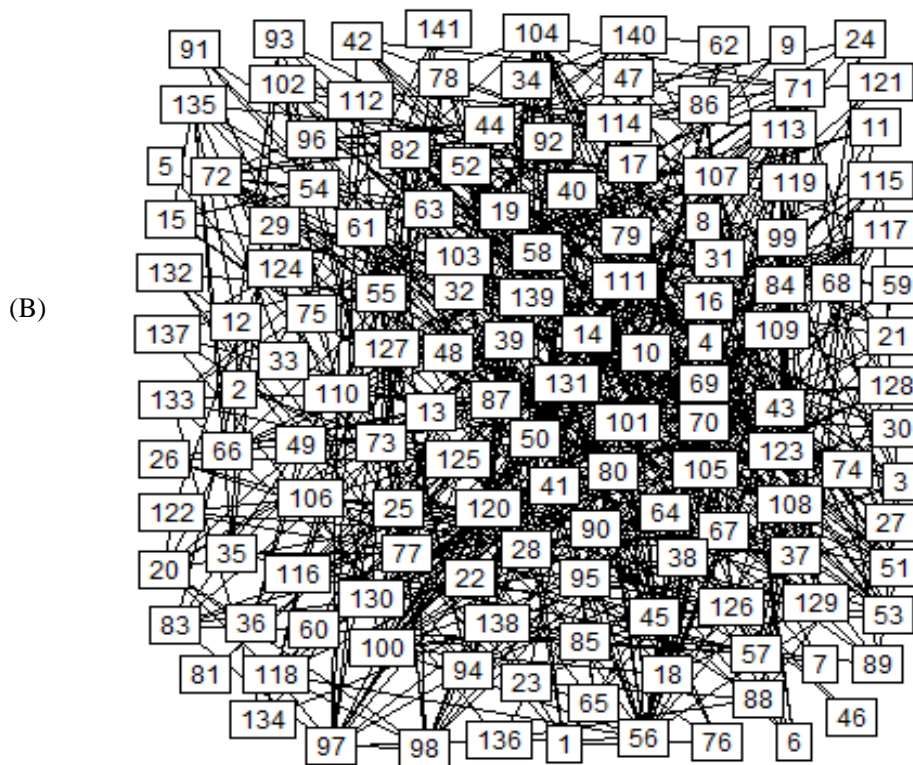


Fig. 2 Between-family (A) and between-species (B) Pearson linear correlations (PH-Apr)

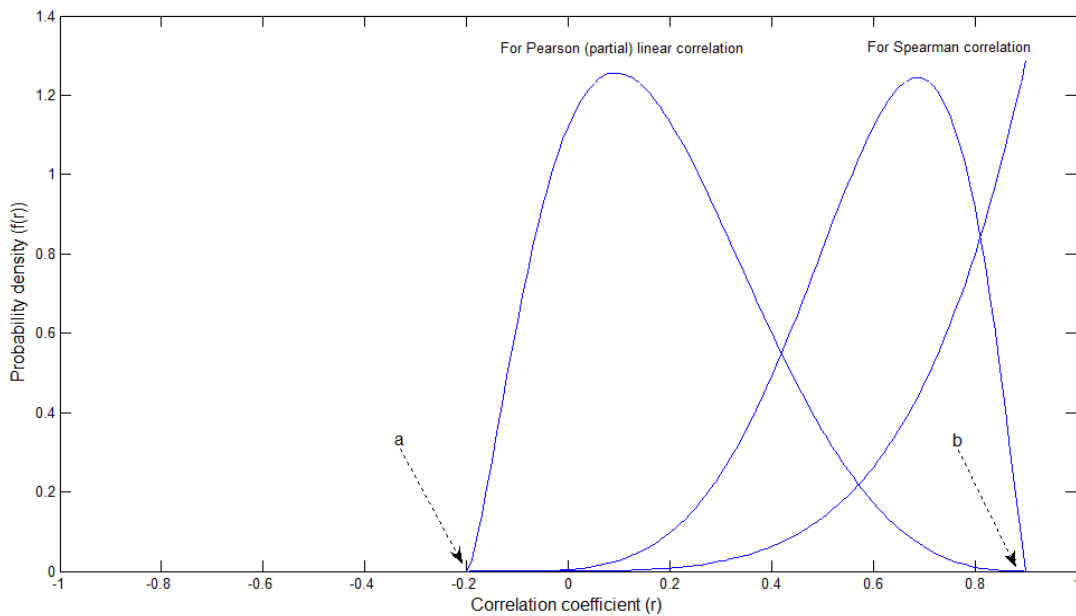


Fig. 3 A distribution model developed by author

The general statistics of ecological networks of arthropods based on Pearson linear correlation and partial linear correlation are listed in Table 1. From Table 1 we may find that both mean partial linear correlation and mean Pearson linear correlation for all levels of taxa and all sampling dates and sites are nearly 0.

The result, obtained from the data in Table 1, showed that the proportion of positive linear (partial linear) correlations declined significantly with the increase in taxa (species, family, (macro) functional group) in the ecosystem (Fig. 4):

$$y=0.6667-0.0033S, r^2=0.467, F=21.921, n=27, p=0.0001 \quad (\text{linear})$$

$$y=0.6916-0.0091S, r^2=0.505, F=13.179, n=15, p=0.003 \quad (\text{partial linear})$$

where y : proportion of positive linear (partial linear) correlations; S : number of taxa. The addition of a taxon will result in a decrease of proportion of positive (partial linear) correlations in the system by 0.33% (0.91%). This result shows that pure linear correlation is stricter than Pearson linear correlation, and thus declined more quickly.

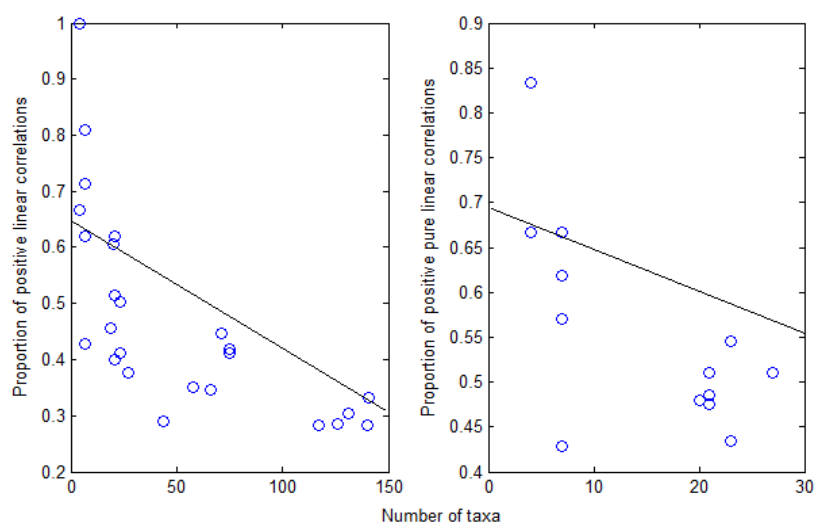


Fig. 4 The relationship between the proportion of positive linear (partial linear) correlations and number of taxa.

The mean linear (partial linear) correlation declined significantly with the increase in taxa (species, family, (macro) functional group) in the ecosystem:

$$y=0.124-0.001S, r^2=0.331, F=11.869, n=27, p=0.0021 \quad (\text{linear})$$

$$y=0.146-0.006S, r^2=0.597, F=19.247, n=15, p=0.0007 \quad (\text{partial linear})$$

where y : mean linear (partial linear) correlation; S : number of taxa. This means that network compactness decreases with the increase of taxa.

4 Degree Distribution and Degree-Taxon Relationships

4.1 Degree distribution

From Table 1 we may find that more than 90% of interactions are positive interactions. The average connectance is 9.8% (9.3%) for (partial) linear correlation generated network. The lower connectance proves that many taxa in the linear network are isolated taxa.

The degree distribution of taxon (species, family, functional group) constructed by between-taxon Pearson

linear and partial linear correlations are overall power law distribution: $L(x)=x^{-\lambda}$, where $L(x)$ is frequency, x is degree, i.e., the number of links of a taxon, as illustrated in Fig. 5.

Table 1 Network information based on Pearson linear correlation and partial linear correlation

Data set	Taxon	Pearson linear corr. (PLC)	Sample size	No. taxa (S)	Total correlations				Number of statistically significant interactions ($p<=0.01$)			
					Total No. (N)	Posi. (s)	s/N (%)	Mean of PLC	Total (L)	L/N (%)	Posi. interact. (w)	w/n (%)
CN-06sep	Func. group	PLC	35	4	6	6	100	0.2791	2	33.3	2	100
CN-06sep	Func. group	Partial PLC	35	4	6	4	66.7	0.2032	1	16.7	1	100
CN-06sep	Func. group	PLC	54	4	6	6	100	0.2663	1	16.7	1	100
CN-06sep	Func. group	Partial PLC	54	4	6	5	83.3	0.1804	2	33.3	2	100
CN-06Oct	Func. group	PLC	60	4	6	4	66.7	0.1698	2	33.3	2	100
CN-06Oct	Func. group	Partial PLC	60	4	6	4	66.7	0.1285	2	33.3	2	100
CN-06Oct	Func. group	PLC	60	4	6	6	100	0.1791	0	0	0	-
CN-06Oct	Func. group	Partial PLC	60	4	6	4	66.7	0.1364	0	0	0	-
PH-Mar	Func. group	PLC	60	21	210	84	40.0	-0.0003	9	4.3	5	55.6
PH-Mar	Func. group	Partial PLC	60	21	210	100	47.6	0.0029	3	1.4	2	66.7
PH-Apr	Func. group	PLC	60	20	190	115	60.5	0.0977	22	11.6	22	100
PH-Apr	Func. group	Partial PLC	60	20	190	91	47.9	0.0151	12	6.3	10	83.3
PH-Sep	Func. group	PLC	60	21	210	130	61.9	0.0766	13	6.2	13	100
PH-Sep	Func. group	Partial PLC	60	21	210	107	51.0	0.0259	4	1.9	4	100
PH-Oct	Func. group	PLC	60	21	210	108	51.4	0.0416	11	5.2	11	100
PH-Oct	Func. group	Partial PLC	60	21	210	102	48.6	0.0178	3	1.4	3	100
PH-Mar	Macro func. group	PLC	60	7	21	9	42.9	-0.0045	1	4.8	1	100
PH-Mar	Macro func. group	Partial PLC	60	7	21	9	42.9	-0.0012	2	9.5	1	50.0
PH-Apr	Macro func. group	PLC	60	7	21	17	81.0	0.2427	7	33.3	7	100
PH-Apr	Macro func. group	Partial PLC	60	7	21	13	61.9	0.0952	2	9.5	2	100
PH-Sep	Macro func. group	PLC	60	7	21	15	71.4	0.169	6	28.6	6	100
PH-Sep	Macro func. group	Partial PLC	60	7	21	14	66.7	0.0963	1	4.8	1	100
PH-Oct	Macro func. group	PLC	60	7	21	13	61.9	0.0757	2	9.5	2	100
PH-Oct	Macro func. group	Partial PLC	60	7	21	12	57.1	0.0508	2	9.5	2	100
CN-06Sep	Family	PLC	35	19	171	78	45.6	0.0557	10	5.9	10	100
CN-06Sep	Family	PLC	54	23	253	127	50.2	0.0529	19	7.5	19	100
CN-06Sep	Family	Partial PLC	54	23	253	138	54.6	0.0276	12	4.7	10	83.3
CN-06Oct	Family	PLC	60	23	253	104	41.1	0.0374	16	6.3	16	100
CN-06Oct	Family	Partial PLC	60	23	253	110	43.5	0.0082	7	2.8	7	100
CN-06Oct	Family	PLC	60	27	351	132	37.6	0.032	24	6.8	23	95.8
CN-06Oct	Family	Partial PLC	60	27	351	179	51.0	0.0171	15	4.3	12	80.0
CN-08Apr	Family	PLC	55	44	946	275	29.1	0.0339	66	7.0	66	100
CN-08Sep	Family	PLC	80	58	1653	582	35.2	-	61	3.7	61	100
PH-Mar	Family	PLC	60	66	2145	742	34.6	-0.001	65	3.0	62	95.4
PH-Apr	Family	PLC	60	71	2485	1107	44.6	0.0376	160	6.4	159	99.4
PH-Sep	Family	PLC	60	75	2775	1161	41.8	0.0359	144	5.2	144	100
PH-Oct	Family	PLC	60	75	2775	1140	41.1	0.0258	117	4.2	116	99.2
PH-Mar	Species	PLC	60	126	7875	2252	28.6	0.0014	272	3.5	264	97.1
PH-Apr	Species	PLC	60	141	9870	3264	33.1	0.0146	572	5.8	559	97.7
PH-Sep	Species	PLC	60	131	8515	2584	30.4	0.0197	411	4.8	410	99.8
PH-Oct	Species	PLC	60	140	9730	2766	28.4	0.012	397	4.1	396	99.8
CN-08Sep	Species	PLC	80	117	6786	1913	28.2	0.0068	256	3.8	255	99.6

Table 2 demonstrates that overall the parameter λ in power law distribution, $L(x)=x^{-\lambda}$, increases as the decline of taxon hierarchy for partial linear correlation based network. For macro functional group, functional group and family, the parameter λ is 0.32, 1.18 and 1.39 (Functional group of China is equivalent to macro

functional group in Philippines). This is also in accordant with the findings that degree centralization changes both qualitatively and quantitatively with size (Butts, 2006). We can conclude that parameter λ in power law distribution, $L(x)=x^{-\lambda}$, is in average 0.8.

The result also means that parameter λ in power law distribution may be scale-dependent.

Table 2 Parameter λ in power law distribution, $L(x)=x^{-\lambda}$.

Data set	CN-06oct	CN-06sept	CN-06oct	PH-Mar	PH-Oct	PH-Oct	
Taxon	Family	Family	Family	Func. group	Func. group	Func. group	
λ	1.6397	1.4307	1.1139	1.2041	1.2041	1.1461	Mean
Data set	CN-06sep	CN-06sep	PH-Apr	PH-Mar	PH-Oct	PH-Sep	0.7933
Taxon	Func. group	Func. group	Macro func. group	Macro func. group	Macro func. group	Macro func. group	
λ	0.2314	0	0.6021	0.6021	0.6021	0.5372	

Note: Networks were created by partial linear correlation.

4.2 Degree-taxon relationships

Table 3 shows that both the number of interactions and the number of positive interactions in the network increase with the number of taxa (species, family, (macro) functional group) for both linear and partial linear correlations based networks. The addition of a taxon will result in an increase of 0.4 (0.3) interactions (positive interactions) in the partial linear correlation based network. And the addition of a taxon will result in an increase of 3 interactions (positive interactions) in the linear correlation based network. However, mean number of (positive) interactions per taxon will not change as the increase of taxa in the partial linear correlation network.

From Table 3, for partial linear correlation based network, the network connectance decreases with the number of taxa ($p < 0.05$). The increase of 10 taxa will result in the 0.07 decrease of connectance.

Two functional group networks, with several isolated taxa, were created by partial linear correlation and Pearson linear correlation respectively, as illustrated in Fig. 6. From Fig. 6, we may find that pollen feeder, external plant feeder, and collector (gatherer, deposit feeder), etc., are key functional groups.

It was found that network structure changed with season and location. The same taxon in the network would connect to different taxa as the change of season and location. A higher level of species aggregation would result in a more stable network structure. Construction of the specific network for specific season and location is thus necessary.

Table 3 Degree-taxon linear regression relationships ($y=a+bS$, S : number of taxa).

	y	a	b	r^2	F	p
Linear correlation based networks	Number of interactions	-43.677	3.045	0.877	178.368	0.000
	Number of positive interactions	-43.224	3.008	0.879	180.946	0.000
	Mean number of interactions per taxon	0.289	0.019	0.856	148.627	0.000
	Mean number of positive interactions per taxon	0.289	0.019	0.850	142.104	0.000
Partial linear correlation based networks	Number of interactions	-0.957	0.412	0.604	19.830	0.001
	Network connectance	0.187	-0.007	0.346	6.873	0.021
	Number of positive interactions	-0.651	0.344	0.624	21.593	0.001
	Mean number of interactions per taxon	0.251	0.005	0.054	0.739	0.406
	Mean number of positive interactions per taxon	0.248	0.003	0.018	0.241	0.632

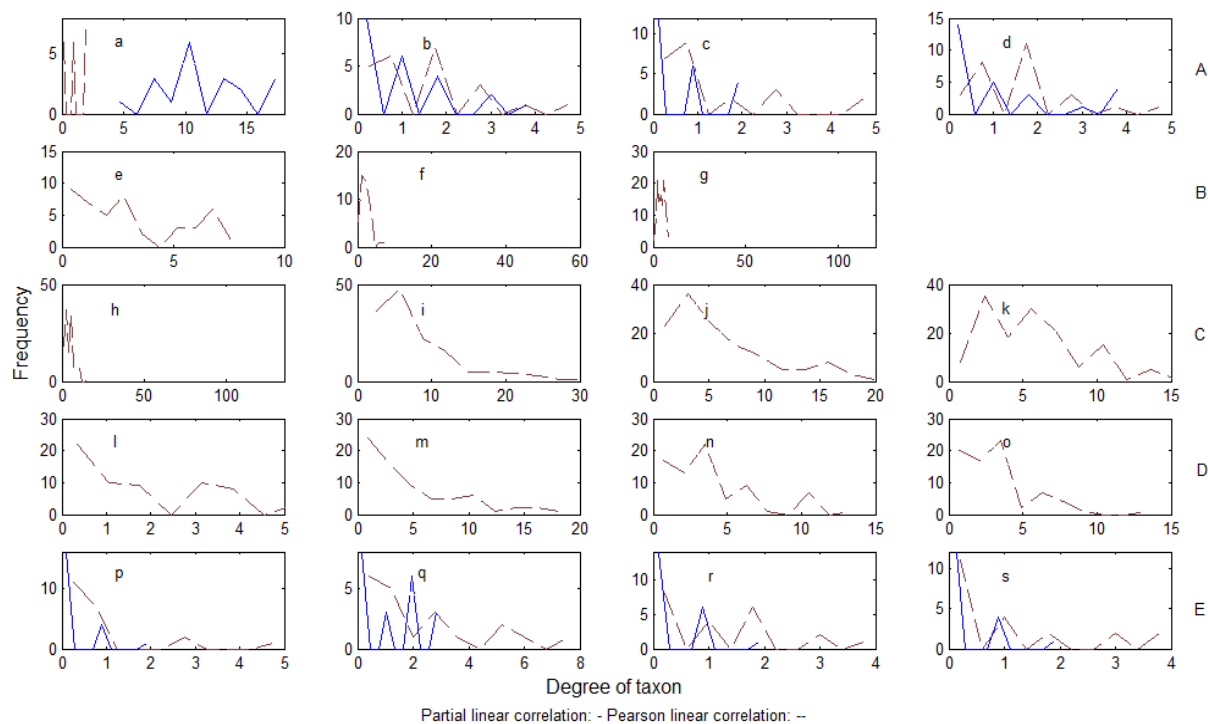
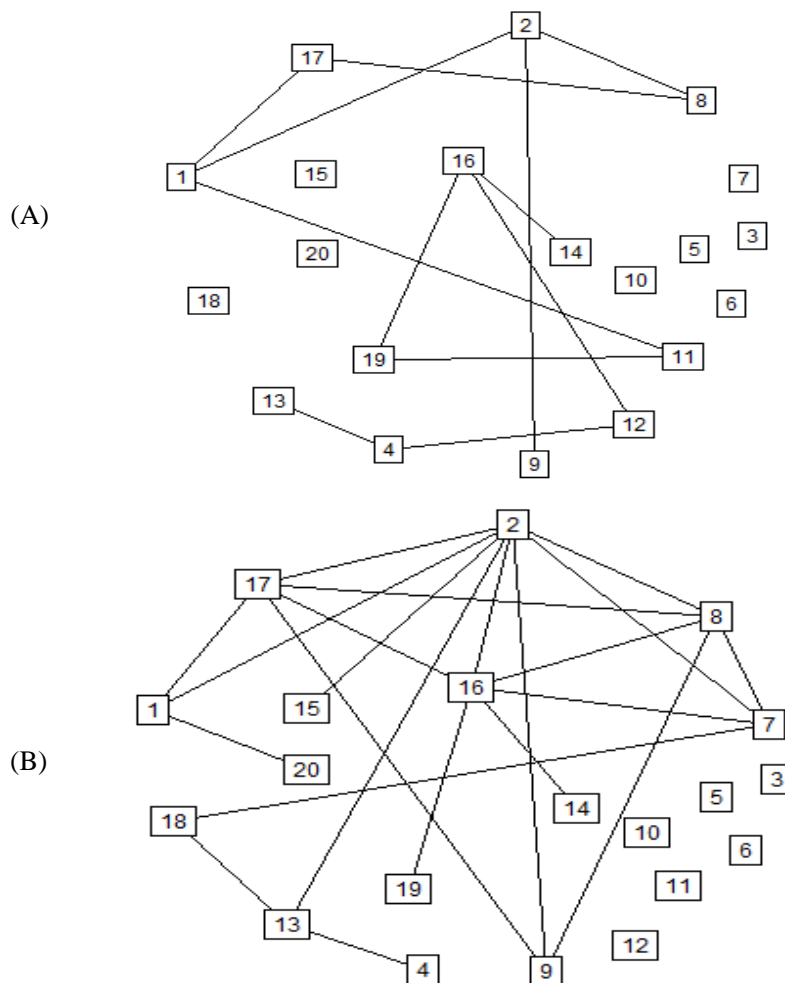


Fig. 5 Degree distributions of ecological networks. A: distributions of family networks for four seasons, China, 2006; B: distributions of family networks for April and September and species networks for September, China, 2008; C, D and E: distributions of species, family and functional group networks for four seasons, Philippines.



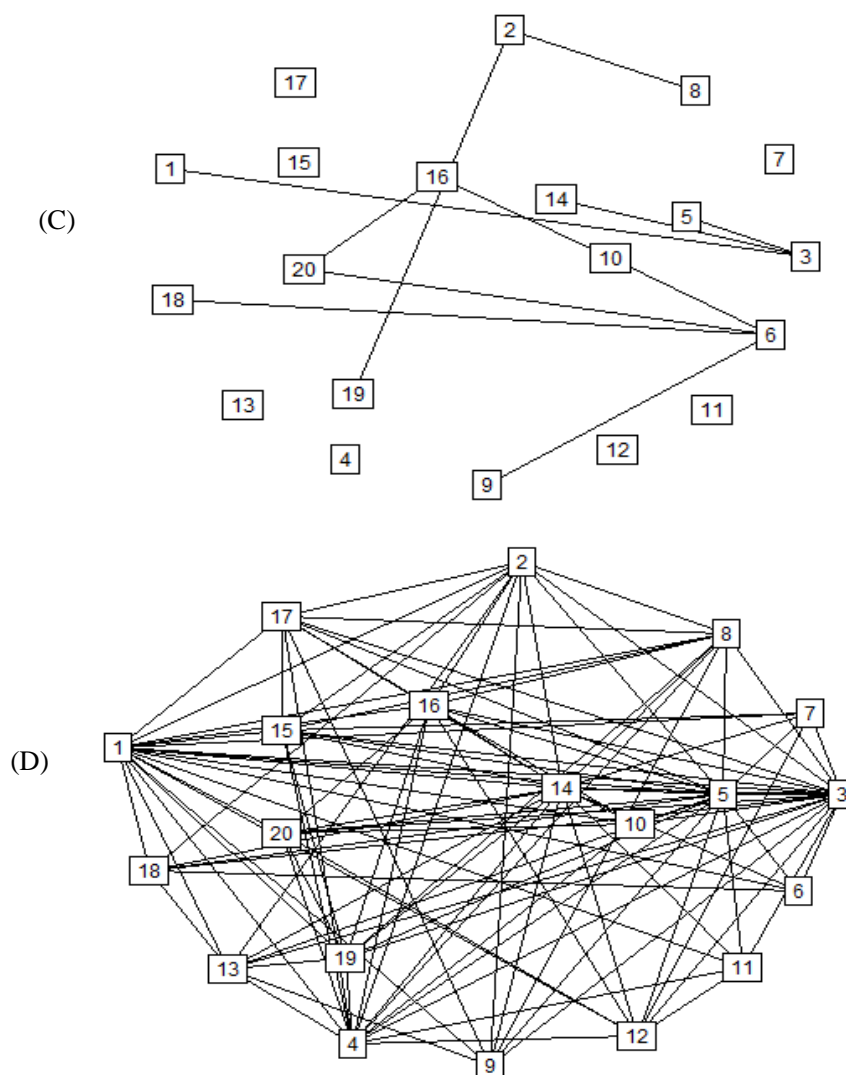


Fig. 6 Functional group networks (PH-Apr) generated by (A) partial linear correlation; (B) Pearson linear correlation; (C): Spearman partial correlation; (D) Spearman correlation. All networks were drawn by the software of Zhang (2012). Functional groups 1 to 20: (1) pollen feeder; (2) external plant feeder; (3) leaf roller/webber; (4) leaf miner; (5) gall former, (6) mixed (Schoenly and Zhang, 1999a); (7) terrestrial flyer; (8) terrestrial crawler, walker, jumper, or hunter; (9) neustonic (water surface) swimmer (semiaquatic); (10) planktonic (water column) swimmer and diver; (11) terrestrial web-builder; (12) terrestrial blood sucker; (13) flying adult that is searching, ovipositing, or larvipositing; (14) idiobiont (acarine ectoparasitoid); (15) collector (filterer, suspension feeder); (16) collector (gatherer, deposit feeder); (17) shredder, chewer of coarse particulate matter; (18) tourist (nonpredatory species with no known functional role other than as prey in ecosystem); (19) herbivore, predator, and detritivore; (20) predator and parasitoid.

5 Quasi-linear Network

General statistics of ecological networks of arthropods based on Spearman correlation are listed in Table 4.

From Table 4 we may find that mean Spearman (partial) correlation (Spearman partial correlation was defined following the calculation of partial linear correlation, which is an approximation of true Spearman partial correlation. However, the formulae for true Spearman partial correlation should be further studied) for all levels of taxa and all sampling dates and sites are nearly 0.5 (0.1). About 99% (91%) between-taxon interactions are positive interactions.

Results showed that the number of positive Spearman correlations and mean Spearman correlation increased with the number of taxa. An ecosystem with more taxa will have a higher mean Spearman correlation.

The number of (positive) interactions and proportion of positive interactions increase with the increase in taxa. The addition of a taxon will result in an increase of 0.4 (0.3) interactions (positive interactions) in the network, which is similar to the results from linear networks.

The average connectance of quasi-linear networks is 15.4% for Spearman partial correlation (but 69.5% for Spearman correlation) based network, larger than the connectance (9.3% and 9.8%) of linear networks.

From Table 3, for Spearman partial correlation based network, the network connectance decreases with the number of taxa ($p < 0.05$). The increase of 10 taxa will result in the 0.1 decrease of connectance.

Table 4 Network information based on Spearman correlation

Data set	Taxon	Spearman correlation coefficient (SCC)	Sampl. size	No. taxa (S)	Total potential interactions				Number of statistically significant interactions ($p < 0.01$)			
					Total No. (N)	Posi. (s)	s/N (%)	Mean of SCC	Total (L)	L/N (%)	Posi. interact. (w)	w/n (%)
CN-06sep	Func. group	SCC	35	4	6	6	100	0.4178	3	50.0	3	100
CN-06sep	Func. group	Parti. SCC	35	4	6	5	83.3	0.2452	2	33.3	2	100
CN-06sep	Func. group	SCC	54	4	6	6	100	0.4058	5	83.3	5	100
CN-06sep	Func. group	Parti. SCC	54	4	6	5	83.3	0.2299	2	33.3	2	100
CN-06oct	Func. group	SCC	60	4	6	6	100	0.3309	2	33.3	2	100
CN-06oct	Func. group	Parti. SCC	60	4	6	5	83.3	0.1986	2	33.3	2	100
CN-06oct	Func. group	SCC	60	4	6	6	100	0.2551	1	16.7	1	100
CN-06oct	Func. group	Parti. SCC	60	4	6	5	83.3	0.1754	1	16.7	1	100
PH-Mar	Func. group	SCC	60	21	210	177	84.3	0.2747	98	46.7	95	96.9
PH-Mar	Func. group	Parti. SCC	60	21	210	132	62.9	0.0555	10	4.8	8	80.0
PH-Apr	Func. group	SCC	60	20	190	181	95.3	0.3749	111	58.4	111	100
PH-Apr	Func. group	Parti. SCC	60	20	190	117	61.6	0.0617	11	5.8	10	90.9
PH-Sep	Func. group	SCC	60	21	210	200	95.2	0.3896	126	60.0	125	99.2
PH-Sep	Func. group	Parti. SCC	60	21	210	131	62.4	0.0636	8	3.8	7	87.5
PH-Oct	Func. group	SCC	60	21	210	208	99.1	0.399	129	61.4	129	100
PH-Oct	Func. group	Parti. SCC	60	21	210	135	64.3	0.0597	9	4.3	7	77.8
PH-Mar	Macro func. group	SCC	60	7	21	15	71.4	0.1541	7	33.3	7	100
PH-Apr	Macro func. group	SCC	60	7	21	20	95.2	0.321	9	42.9	9	100
PH-Sep	Macro Func. group	SCC	60	7	21	17	81.0	0.2152	7	33.3	6	85.7
PH-Oct	Macro func. group	SCC	60	7	21	20	95.2	0.2688	8	38.1	8	100
CN-06sep	Family	SCC	35	19	171	171	100	0.5534	123	71.9	123	100
CN-06sep	Family	SCC	54	23	253	251	99.2	0.5732	218	86.2	218	100
CN-06oct	Family	SCC	60	23	253	253	100	0.5769	219	86.6	219	100
CN-06oct	Family	SCC	60	27	351	349	99.4	0.6035	306	87.2	306	100
CN-06oct	Family	Parti. SCC	60	27	351	205	58.4	0.0455	11	3.1	9	81.8
CN-08apr	Family	SCC	55	44	946	946	100	0.7252	914	96.6	914	100
CN-08Sep	Family	SCC	80	58	1653	1630	98.6	0.5625	1432	86.6	1432	100
PH-Mar	Family	SCC	60	66	2145	2077	96.8	0.5191	1683	78.5	1680	99.8
PH-Apr	Family	SCC	60	71	2485	2446	98.4	0.5311	2035	81.9	2034	100
PH-Sep	Family	SCC	60	75	2775	2756	99.3	0.6009	2467	88.9	2466	100
PH-Oct	Family	SCC	60	75	2775	2763	99.6	0.5792	2404	86.6	2404	100
PH-Mar	Species	SCC	60	126	7875	7785	98.9	0.6391	7067	89.7	7064	100
PH-Apr	Species	SCC	60	141	9870	9805	99.3	0.6341	9044	91.6	9042	100
PH-Sep	Species	SCC	60	131	8515	8494	99.8	0.7054	8134	95.5	8133	100
PH-Oct	Species	SCC	60	140	9730	9713	99.8	0.6995	9268	95.3	9268	100
CN-08sep	Species	SCC	80	117	6786	6769	99.8	0.703	6555	96.6	6555	100

The degree distribution of taxon derived from between-taxon Spearman partial correlation is power law distribution, $L(x)=x^{-\lambda}$, where x is degree, i.e., the number of links of a taxon, as illustrated in Fig. 4. It means that the link frequency of the quasi-linear network decreases as the degree. We can conclude that parameter λ in power law distribution, $L(x)=x^{-\lambda}$, is in average 0.86.

According to parameter λ in power law distribution, $L(x)=x^{-\lambda}$, the degree distributions of networks constructed by partial linear correlation and Spearman partial correlation are similar to each other.

Table 5 Degree-taxa linear regression relationships ($y=a+bS$, S : number of taxa) based on Spearman correlation and quasi-linear networks.

	y	a	b	r^2	F	p
Spearman correlation	Number of positive correlations	-1042.656	67.476	0.937	372.505	0
	Proportion of positive correlations	0.944	0.0005	0.101	2.796	0.107
	Mean correlation	0.356	0.0027	0.571	33.299	0
Spearman correlation	Number of interactions	-2010.660	63.288	0.927	317.842	0
	Number of positive interactions	-1020.863	63.279	0.927	317.622	0
	Mean number of interactions per taxon	-2.514	0.478	0.995	4684.603	0
	Mean number of positive interactions per taxon	-2.535	0.478	0.995	4592.662	0
Spearman partial correlation	Number of interactions	0.116	0.436	0.944	117.932	0
	Network connectance	0.340	-0.013	0.844	37.759	0.001
	Number of positive interactions	0.474	0.347	0.906	67.361	0
	Proportion of positive interactions	1.034	-0.009	0.843	37.665	0.001
	Mean number of interactions per taxon	0.442	0.0001	0.0002	0.001	0.972
	Mean number of positive interactions per taxon	0.456	-0.004	0.140	1.142	0.321

Table 6 Parameter λ in power law distribution, $L(x)=x^{-\lambda}$.

Data set	CN-06sep	PH-Mar	CN-06oct	PH-Sep	PH-Oct	PH-Apr	
Taxon	Family	Func. group	Family	Func. group	Func. group	Func. group	
λ	3.2181	1.3652	1.3098	1.2640	1.2137	1.2091	Mean
Data set	PH-Apr	CN-06oct	CN-06sep	CN-06oct	PH-Mar	PH-Oct	0.8574
Taxon	Macro func. group	Func. group	Func. group	Func. group	Macro func. group	Macro func. group	
λ	0.4771	0.2314	0	0	0	0	

Note: Networks were constructed by Spearman partial correlation.

6 Discussion

The scale-free property is an important feature of complex networks. Researchers are trying to approach the causes and mechanisms of the property in recent years. The possible causes and mechanisms include the following: (1) with the addition of new vertices, the network continues to expand; (2) new vertices tend to connect to already better connected vertices (Barabasi and Albert, 1999; Barabasi, 2009).

In present study, I defined the linear network and the quasi-linear network. Quasi-linear networks can be constructed based on other quasi-linear correlations, as point correlation, linkage correlation, etc. (Zhang, 2007b). However, the partial correlation based on these correlation measures should be further defined, as temporarily defined and used for Spearman partial correlation in present study. To obtain a relatively stable network, all sampling plots or sites should first be homogeneous, i.e., all plots are the same in the environmental conditions. It should be noted that sample size (number of samples) must be larger than number of taxa in the community sampling in order to better use partial correlation analysis.

It was found that network structure changed with season and location. The same taxon in the network would connect to different taxa as the change of season and location. A higher level of species aggregation

would result in a more stable network structure. Construction of the specific network for specific season and location is necessary.

Studies on between-taxon interactions may encounter problems due to the fluidity of the interacting units, and the fact that sub-taxa of a larger taxon may themselves interact with others both within and without the “parent” (Butts, 2009). To avoid mistakes, the taxa (species aggregates) must be deliberately defined. Large sample sets should also be taken.

Most of the past studies stressed the importance of negative interactions, as competitive exclusion principle, Lotka-Volterra model, etc. Those theories have laid a foundation for theoretical framework of ecology (Bruno et al., 2003). However, more and more recent studies have challenged the dominance of competition (Feinsinger, 1987; Callaway, 1995; Bruno et al., 2003; Dormann, 2011). Some argued that mutualism is the basis of community process (Bengtsson et al. 1994; Palmer et al., 2003). Mutualistic networks included mutually beneficial interactions (positive interactions) which play a major role in the generation and maintenance of biodiversity on Earth (Thompson, 2005). And mutualistic interactions are always weak (Bascompte and Jordano, 2007). A report showed that there were 14 positive interactions in 17 interactions of the pollination network (Hegland et al., 2009). To explain mutualism some hypotheses have been developed. Neutral hypothesis argues that network patterns are generated by random interactions of pair-wise individuals which lead to more frequent interactions between abundant species than between rare species and more species interacts with each other (Dupont et al., 2003, Ollerton et al., 2003, Vazquez et al., 2007). Link banning hypothesis maintains that network patterns are jointly generated by species phenotype, bio-climate, spatial distribution and phelogeny (Jordano et al., 2003; Rezende et al., 2007; Santamara and Rodriguez-Girones, 2007; Stang et al., 2007). The constant connectance hypothesis reported in some previous studies (Pimm et al., 1991; Havens, 1992; Martinez, 1992) did not hold for the networks we investigated. Present study proved that the parameter λ in power law distribution of degree, $p(x)=x^{-\lambda}$, increases as the decline of taxon hierarchy and λ was in average 0.8 to 0.9. Network connectance decreases with the number of taxa ($p<0.05$). The addition of 10 taxa results in the 0.07 to 0.1 decrease of connectance.

Link-species scaling law supposed that $L \approx aS$, on average, the number of links per species in a web is constant and scale invariant at roughly two, and therefore, $L \approx 2S$ (Cohen et al., 1990; Martinez, 1992). However, the present results showed that the number of links per taxon was 0.25 for linear networks and 0.45 for quail-linear networks.

To better represent the importance of a taxon (as a producer) in the network, the degree is a reasonable measure. In general the more the degree of a taxon, the more significant the taxon (as a producer) is.

The present study proved that most correlations are weak, and positive interactions accounted for the most of the actual interactions. In my view, the sampling data for mutualistic species and predator/prey and parasite(parasitoid)/host species will in most cases result in positive correlations, as theoretically illustrated in Fig. 7. The sampling data for only a few of competitive species (in particular the species with perfect competition) would yield negative correlations. Dominance of negative interactions may occur in competition driven communities. Overall positive interactions, caused mainly by mutualism, predation/parasitism, etc., are thus dominant in the ecological network. The dominance of positive interactions is a natural selection for co-existence of species and is the outcome of ecosystem and species evolution. I suggest that the number and portion of positive interactions may be the most important indices for community stability and functionality. Mutualism is the most significant trophic relationship, seconded by predation/parasitism, competition is the worst for community stability.

Similar to most studies of ecological networks, this study focused on the statistical aspect of networks. So far most of the studies have not yet used graph theory and optimization theory to analyze ecological networks

(Ferrarini, 2011; Zhang, 2011a; Zhang, 2012). Apart from statistical analysis and mechanism exploitation of ecological networks (Montoya and Sole, 2003; Cohen and Briand, 1984; Cohen and Newman, 1988), future works should also be put on these aspects: (1) construction of practical ecological networks; (2) analysis and application of graphic properties of networks; (3) dynamic modeling of network structure and functionality and mechanism exploitation of network dynamics using agent-based modeling (Zhang, 2012); (4) design and optimization of networks.

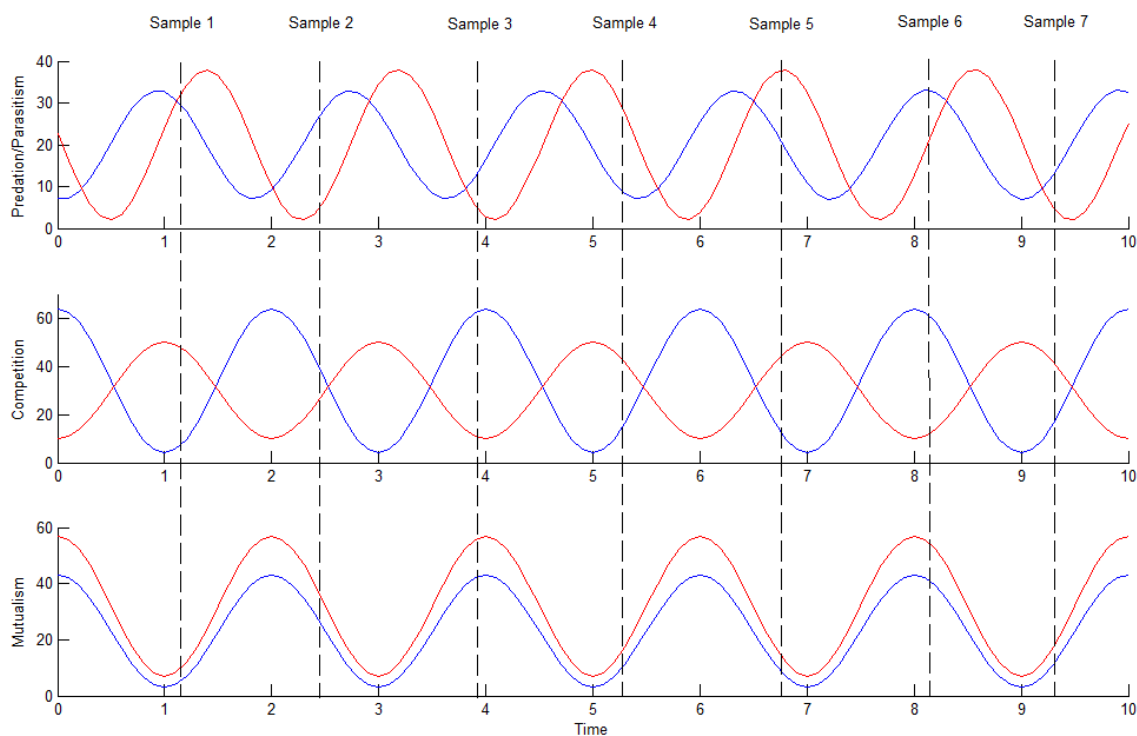


Fig. 7 Positive correlations for predation/parasitism and mutualism and negative correlation for some competition. Predation/Parasitism curves represent Lotka-Volterra model. Competition curves represent perfect competition.

Acknowledgement

The author thanks all field investigators of rice arthropods both in China and Philippines.

References

- Arii K, Derome R, Parrott L. 2007. Examining the potential effects of species aggregation on the network structure of food webs. *Bulletin of Mathematical Biology*, 69: 119-133
- Barabasi A, Albert R. 1999. Emergence of scaling in random networks. *Science*, 286: 509-512
- Barabasi AL. 2009. Scale-free networks: a decade and beyond. *Science*, 325: 412-413
- Bascompte J. 2009. Disentangling the web of life. *Science*, 325: 416-419
- Bascompte J, Jordano P. 2007. The structure of plant-animal mutualistic networks: the architecture of biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 38: 567-593
- Bengtsson G, Hedlund K, Rundgren S. 1994. Food and density related dispersal in a soil Collembola. *Journal of Animal Ecology*, 63: 513-520
- Bengtsson G, Erlandsson A, Rundgren S. 1988. Fungal odour attracts soil Collembola. *Soil Biology and Biochemistry*, 20: 25-30

- Bruno JF, Stachowicz JJ, Bertness MD. 2003. Inclusion of facilitation into ecological theory. *Trends in Ecology and Evolution*, 18(3): 119-125
- Butts CT. 2006. Exact bounds for degree centralization. *Social Networks*, 28: 283-296
- Butts CT. 2009. Revisiting the foundations of network analysis. *Science*, 325: 414-416
- Callaway RM. 1995. Positive interactions among plants (Interpreting botanical progress). *The Botanical Review*, 61: 306-349
- Cohen JE, Briand. 1984. Trophic links of community food web. *Proceedings of the National Academy of Sciences of the USA*, 81: 4105-4109
- Cohen JE, Briand F, Newman CM. 1990. *Community Food Webs: Data and Theory*. Biomathematics (Vol.20), Springer, Berlin, Germany
- Cohen JE, Newman CM. 1988. Dynamic basis of food web. *Ecology*, 69: 1655-1664
- Comita LS, Muller-Landau HC, Aguilar S, et al. 2010. Asymmetric density dependence shapes species abundances in a tropical tree community. *Science*, 329: 330-332
- Dormann CF. 2011. How to be a specialist? Quantifying specialisation in pollination networks. *Network Biology*, 1(1): 1-20
- Dunne JA, Williams R, Martinez N. 2002. Conservation of species interaction networks. *Ecology Letters*, 5: 558
- Dupont YL, Hansen DM, Olesen JM. 2003. Structure of a plant-flower-visitor network in the high altitude sub-alpine desert of Tenerife, Canary Islands. *Ecography*, 26: 301-310
- Faust K. 2007. Very local structure in social networks. *Sociological Methodology*, 37: 209-256
- Feinsinger P. 1987. Professional ecologists and the education of young children (Commentary). *Trends in Ecology and Evolution*, 2:51-52
- Ferrarini A. 2011. Some steps forward in semi-quantitative networks modelling. *Network Biology*, 1(1):72-78
- Havens KE, Bull LA, Warren GL. 1996. Food web structure in a subtropical lake ecosystem. *Oikos*, 75(1): 20-32
- Hegland SJ, Grytnes JA, Totland O. 2009. The relative importance of positive and negative interactions for pollinator attraction in a plant community. *Ecological Research*, 24: 929-936
- Martinez ND. 1992. Constant connectance in community food webs. *American Naturalist*, 139: 1208-1218
- Martinez ND. 1993a. Effects of scale on food web structure. *Science*, 260: 242-243
- Martinez ND. 1993b. Effects of resolution on food web structure. *Oikos*, 66: 403-412
- Matsuda H, Namba T. 1991. Food web graph of a coevolutionarily stable community. *Ecology*, 72: 267-276
- May RM. 1983. The structure of food webs. *Nature*, 301: 566-568
- Memmott J, Craze PG, Waser NM. 2007. Global warming and the disruption of plant-pollinator interactions. *Ecology Letters*, 10: 710-717
- Montoya JM, Pimm SL, Sole RV. 2006. Ecological networks and their fragility. *Nature*, 442: 259-264
- Montoya JM, Sole RV. 2003. Topological properties of food webs: from real data to community assembly models. *Oikos*, 102: 614-622
- Ollerton J, Johnson SD, Cranmer L, et al. 2003. The pollination ecology of an assemblage of grassland asclepiads in South Africa. *Annals of Botany*, 92: 807-834
- Onnela JP, et al. 2007. Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences of USA*, 104: 7332
- Paine RT. 1980. Food webs: linkage, interaction strength and community infrastructure. *Journal of Animal Ecology*, 49: 667-686
- Paine RT. 1988. Food webs: road maps of interactions or grist for theoretical development? *Ecology*, 69:

1648–1654

- Paine RT. 1992. Food-web analysis through field measurement of per capita interaction strength. *Nature*, 355: 73-75
- Palmer TM, Stanton ML, Young TP. 2003. Competition and coexistence: exploring mechanisms that restrict and maintain diversity within mutualist guilds. *The American Naturalist*, 162: S63-S79
- Petchey OL, Beckerman AP, Riede JO, et al. 2008. Size, foraging, and food web structure. *Proceedings of the National Academy of Sciences of USA*, 105(11): 4191-4196
- Pimm SL. 1991. *The balance of nature? Ecological issues in the conservation of species and communities.* University of Chicago Press, USA
- Rezende EL, Lavabre JE, Guimaraes PR, et al. 2007. Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature*, 448: 925-928
- Santamara L, Rodriguez-Girones MA. 2007. Linkage Rules for Plant–Pollinator Networks: Trait Complementarity or Exploitation Barriers? *PLoS Biology*, 5(2): e31
- Schmitz OJ, Booth G. 1997. Modelling food web complexity: The consequences of individual-based, spatially explicit behavioural ecology on trophic interactions. *Evolutionary Ecology*, 11: 379-398
- Schoener TW. 1993. On the relative importance of direct versus indirect effects in ecological communities. In: *Mutualism and community organization* (Kawanabe H, Cohen JE and Iwasaki K, eds). Oxford University Press, UK, 365–411
- Schoenly KG, Zhang WJ. 1999a. IRRI Biodiversity Software Series. I. LUMP, LINK, AND JOIN: utility programs for biodiversity research. IRRI Technical Bulletin No. 1. Manila (Philippines): International Rice Research Institute, Manila, Philippines, 1-23
- Schoenly KG, Zhang WJ. 1999b. IRRI Biodiversity Software Series. V. RARE, SPPDISS, and SPPANK: programs for detecting between-sample difference in community structure. IRRI Technical Bulletin No.5. International Rice Research Institute, Manila, Philippines
- Solow AR, Beet A., 1998. On lumping species in food webs. *Ecology*, 79: 2013-2018
- Stang M, Klinkhamer PGL, van der Meijden E. 2007. Asymmetric specialization and extinction risk in plant–flower visitor webs: a matter of morphology or abundance? *Oecologia*, 151: 442-453
- Sugihara G, Bersier LF, Schoenly K. 1997. Effects of taxonomic and trophic aggregation on food web properties. *Oecologia*, 112: 272-284
- Thompson JN. 2005. *The Geographic Mosaic of Coevolution.* University of Chicago Press, Chicago, USA
- Vazquez DP, Melian CJ, Williams NM, et al. 2007. Species abundance and asymmetric interaction strength in ecological networks. *Oikos*, 116: 1120-1127
- Whitehead H, Dufault S. 1999. Techniques for analyzing vertebrate social structure. *Advances in the Study of Behavior*, 28:33-74
- Williams RJ, Martinez ND. 2000. Simple rules yield complex food webs. *Nature*, 404: 180-183
- Zhang WJ. 2007a. Pattern classification and recognition of invertebrate functional groups using self-organizing neural networks. *Environmental Monitoring and Assessment*, 130: 415-422
- Zhang WJ. 2007b. Computer inference of network of ecological interactions from sampling data. *Environmental Monitoring and Assessment*, 124: 253-261
- Zhang WJ. 2011a. Network Biology: an exciting frontier science. *Network Biology*, 1(1):79-80
- Zhang WJ. 2011b. A Java program to test homogeneity of samples and examine sampling completeness. *Network Biology*, 1(2):127-129
- Zhang WJ. 2012. *Computational Ecology: Graphs, Networks and Agent-based Modeling.* World Scientific, Singapore