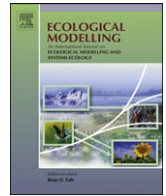




Contents lists available at [SciVerse ScienceDirect](#)

## Ecological Modelling

journal homepage: [www.elsevier.com/locate/ecolmodel](http://www.elsevier.com/locate/ecolmodel)



### A fish tank model for assembling food webs

Wei-Chung Liu<sup>a</sup>, Hsuan-Wien Chen<sup>b,\*</sup>, Tsung-Hsi Tsai<sup>a</sup>, Hsien-Kuei Hwang<sup>a</sup>

<sup>a</sup> Institute of Statistical Science, Academia Sinica, Taiwan

<sup>b</sup> Biodiversity Research Center, Academia Sinica, Taiwan

#### ARTICLE INFO

Article history:  
Available online xxx

Keywords:  
Food web  
Assembly  
Structure  
Trophic link  
Trophic level

#### ABSTRACT

In the past two decades, several models for food webs have been proposed. Those models provide possible explanations on how the food webs we observe today arrive at their respective network structure. Those well established models almost all rely on two common parameters, the numbers of species and trophic links, as well as assuming that species are ranked in some hierarchical order or ecological dimensions. Under those model assumptions species are then allowed to interact with each other depending on their relative positions in relevant ecological dimensions, resulting in simulated food webs that can capture the characteristics of their real-life counterparts. In this paper, we propose a simple model for food webs following biological intuitions. Our model differs from others in that we do not have the prior assumption of species being ordered in any ecological dimensions, and we do not assume a particular number of trophic links. Our model here takes two basic parameters, the number of species and the number of resources, and allows the random encounter between species and resources, and as a result forming trophic interactions. One critical assumption in this model is that after a species obtains resources, it will then become new resources for other species to utilise. In essence, the model here constructs food webs by passing resources among different species. Our model predicts the maximum number of trophic links of a food web when given the number of species, and shows that connectance is a non-linear decreasing function of species number. Furthermore, this simple model also demonstrates that the number of trophic levels tends to decrease when resource number is larger, echoing empirical studies from the literature. We further modified our simple model to fit a particular food web data and suggest the possible mechanism underlying the growth of the empirical food web. Lastly, we assessed the performance of our model by fitting it to a wide range of real food webs and compared the result with those derived from other models.

© 2012 Elsevier B.V. All rights reserved.

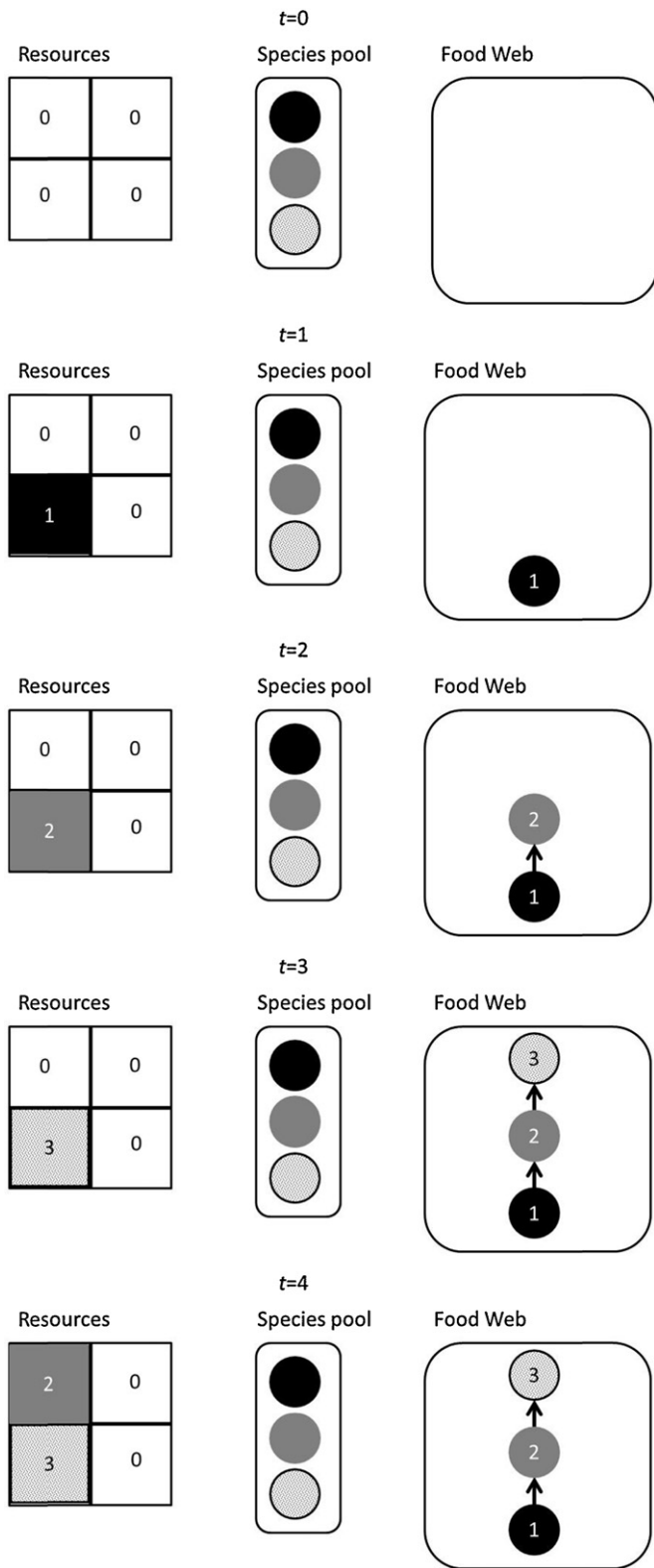
#### 1. Introduction

A food web is a graphical representation of an ecosystem. It basically consists of nodes and links representing species and trophic interactions respectively. The structural organisations of food webs have attracted the attention of ecologists for decades as they are thought to play a role in the stability or the robustness properties of various ecosystems (Pimm et al., 1991; Dunne et al., 2002; Melián and Bascompte, 2002; Estrada, 2007; Dunne and Williams, 2009; Gilbert, 2009). Although food webs possess complex structural properties, yet theoreticians seek for simple rules that give rise to the food webs we observe today (Cohen et al., 1990; Williams and Martinez, 2000; Cattin et al., 2004; Rossberg et al., 2006; Warren et al., 2010); and it is this linkage between simplicity in underlying mechanism to complex macroscopic organisation that poses as one of the major challenges in ecological research (Allesina et al., 2008).

There have been quite a few food models in the literature in the past two decades, with the newer ones incorporating more realistic or complicated rules mimicking mechanisms that seem to operate in nature, and therefore they naturally perform better in fitting real food webs than their older counterparts. The simplest food web model to date is the cascade model where species are assumed to be arranged in a hierarchical order in a linear dimension, with higher species consuming lower ones in a probabilistic manner (Cohen et al., 1990). The interpretation of this hierarchical ordering of species is open to debate, but in general, species attributes such as body size and metabolic rate are potential candidates (Cohen et al., 1993; Woodward et al., 2005). This simple model paves the way for other more complicated ones like the niche model (Williams and Martinez, 2000) where each species is assumed to have a diet range that allows for feeding cycles and cannibalism. Furthermore, the following nested hierarchy model considers a more complex rule of trophic interaction that spans more than a single dimension (i.e. body size, phylogenetic constraints and adaptations) in order to fit the real data better than its predecessors (Cattin et al., 2004).

Those mainstream food web models have at least two things in common. First, there is a prior assumption of species being ordered

\* Corresponding author at: 128 Academia Road, Sec. 2, Biodiversity Research Center, Academia Sinica, Taipei, Taiwan. Tel.: +886 933 419391.  
E-mail address: [chen7@sinica.edu.tw](mailto:chen7@sinica.edu.tw) (H.-W. Chen).



**Fig. 1.** A pictorial representation of the base model. At time  $t=0$ , there is a collection of unoccupied resources, a species pool and an empty food web. At  $t=1$ , the black species is selected to occupy a resource simply by chance, and it now has a trophic level of  $0+1=1$ , and the resource it occupies now has a label value of 1. The present food web only has the black species. At  $t=2$ , a new species (grey species) then enters the system and utilises the resource represented by the black species, and the grey species has a trophic level of  $1+1=2$ , and the resource it now occupies has a label value of 2. The food web now consists of black species and grey species with a trophic link between them. At  $t=3$ , a new species (dotted) is selected to utilise the resource previously occupied by the grey

in one or multiple linear dimensions. Second, they share at least two parameters: one is the number of species in a food web and the other, although only implicitly assumed, is the number of trophic links (although food web models such as the cascade model and the niche model do not explicitly predefine the number of trophic links, but they are constrained by fixed connectance values). Here, we ask if there is a more fundamental model for food webs that does not require the prior assumption of species being ordered in a hierarchical order along a linear dimension, as well as not to fix the number of trophic links a priori. If a layman is asked of how an ecosystem works, then the most probable answer might be something along these lines: plants convert sun light to produce food for herbivores, which then in turn become the food for carnivores, and those carnivores then become the food of other top predators. In this paper, we construct a simple model for food webs based on this simple principle. The paper is organised as follows. First, we describe the simplest version of our model (i.e. the base model), and explore its structural properties. Second, we then fit the base model to a small empirical food web. Third, we propose ways of modifying our base model such that it can portray the basic characteristics of the empirical food web better. Fourth, we then assess the performance of our modified model in fitting several larger datasets and compare our results with those derived from other food web models. Finally, we discuss the implication of our models and their results as well as some notes on future directions.

## 2. Models and results

### 2.1. Base model

Our model is of very simple nature. It assumes that an ecosystem has  $R$  unoccupied resources (taking on integer values) and a pool of  $S$  species waiting to enter an “empty food web”. A resource here is of abstract nature, but it can be thought of as a unit of abiotic entity such as space, land, energy or nutrients. We define  $V_i$  as the label value of resource  $i$  (see below) and set all label values to zero to indicate all resources are unoccupied initially. The purpose of label value of a resource is to indicate what species are eligible to occupy and utilise it (see below). Next we define  $W_j$  as the trophic level of species  $j$ . Here we adopt the concept of trophic level as in Begon et al. (2006): species at trophic level 1 are primary producers (or basal species); those at trophic level 2 are primary consumers; and species at trophic level 3 are secondary consumers, and so on. A species that is yet to appear in the food web has its trophic level undetermined; and a species’ trophic level is determined when it first enters the food web (see below) and remains fixed throughout the entire growth of the food web.

The model is constructed with the following rules. At time  $t$ , one species is sampled from the species pool at random (with replacement). One of the followings can happen:

- (1) If the selected species is a new species (i.e. not present in the food web before time  $t$ ), then it can utilise one resource picked at random from the set of all resources. If resource  $i$  is picked, then the trophic level of this new species equals to  $V_i + 1$ ; and resource  $i$ , now being occupied by this new species, has a new label value equal to the trophic level of this new species.

species. The dotted species has a trophic level of  $2+1=3$ , and the resource it now occupies has a label value of 3. Now the food web consists of black, grey and dotted species with the dotted species consuming grey species, and the grey species consuming black species. Note that there is no trophic link between black species and dotted species because the resource the dotted species utilises was in the form of grey species. At  $t=4$ , the grey species is selected again by random chance, but this time, it can only use resources that have label values lower than its trophic level (i.e. the top-left, top-right and bottom-right resources), and simply by chance it utilises the top-left one. The food web remains unchanged even though the grey species is allowed to enter the system again.

- (2) If the selected species  $j$  is an existing species, then it is allowed to pick a random resource  $i$  that has a label value lower than its trophic level (i.e.  $W_j > V_i$ ). The label value of this particular resource now takes a new value equal to the trophic level of its new owner.
- (3) If the selected species  $j$  is an existing species and if the label value of every resource at time  $t$  is greater than or equal to  $W_j$ , then nothing will happen.

In any cases, a trophic interaction from species  $i$  to species  $j$  occurs if species  $j$  utilises a resource whose most recent owner is species  $i$  (i.e.  $j$  consumes  $i$ ). The occupation and utilisation of a resource here is analogous to a species consumes another species, and itself then becomes a resource for other species at higher trophic levels.

These rules do not necessarily result in a new trophic link being added to the food web at each time step (note that we distinguish trophic links from trophic interaction by referring the former as a directed link between two nodes in a food web). A new trophic link from species  $i$  to species  $j$  is added to the food web if and only if: (1) there was no such a trophic link previously; and (2) species  $j$  occupies a resource that was previously occupied by species  $i$ . The model will not produce a new trophic link if one of the followings occurs: (1) there are no suitable resources for the selected species; (2) the selected species (regardless it is a new one or an existing one) utilises a previously unoccupied resource; and (3) the selected species  $j$  is an existing species and it utilises a resource that was previously occupied by species  $i$ , and there is already a trophic link from species  $i$  to species  $j$  via other resource(s). The simulation of a food web is then terminated when all species are present and when all resources are being occupied by species at the highest trophic level (as no more new trophic links will be added to the food web). Fig. 1 gives a pictorial representation of the base model.

The base model has two parameters: one is the number of species ( $S$ ) and the other is the number of resources ( $R$ ). Note that the number of trophic links here is not a prefixed parameter but is a consequence of the model outcome. We next investigate how two fundamental food web properties, namely the number of trophic links ( $L$ ) and the number of trophic levels ( $T$ ), vary when changing  $R$  and  $S$ . For each parameter combination of  $R$  and  $S$ , we simulated the model 1000 times, and determined the mean number of trophic links and the mean number of trophic levels. For a fixed number of species, the number of trophic links increases initially towards a peak, and then declines to zero when the number of resources approaches infinity (Fig. 2a). For larger species numbers, the curvature is more profound and the number of trophic links peaks at larger values of  $R$  (Fig. 2a).

As we can observe from Fig. 2a, for a given number of species, there is always a maximum number of trophic links attainable by the model. Theoretically speaking, the number of all possible trophic links is  $(S^2 - S)/2$  for a directed graph. When the maximum number of trophic links produced by the model and its theoretical maximum (i.e.  $(S^2 - S)/2$ ) are plotted against the number of species, we can observe that they do not match with each other; in fact, the maximum predicted by the model is far less than its theoretical counterpart (Fig. 3a). To see this more clearly, we can divide the model maximum by the theoretical maximum, thus obtaining the connectance of a model food web, and then plot this against the corresponding number of species; and our result shows that connectance decreases non-linearly with species number (Fig. 3b). Thus, although intuition suggests that increasing species number should increase the number of trophic links, but this does not result in an increase in food web connectance.

As for the number of trophic levels (Fig. 2b), we found that the maximum value is always  $S - 1$  regardless of species number; and this occurs when the number of resources is 1. When the number

of resources increases, the number of trophic levels decreases non-linearly to zero.

## 2.2. Testing the base model with data

Next we fit our base model to a simple food web data. The food web data used here describes the Kuosheng Bay ecosystem in North-East Taiwan (Lin et al., 2004). We use the version modified by Jordán et al. (2009) where there are 15 trophic groups and 36 trophic links (Fig. 4), and in this particular case there are 8 trophic levels. Since we are given 15 species, we only need to systematically vary the number of resources ( $R$ ) and see how the numbers of trophic links and trophic levels change. Again, for each parameter combination, we simulated the model 1000 times, and calculated the mean numbers of trophic links and trophic levels. Fig. 5 shows that the maximum number of trophic links attainable by the model is around 27 and this occurs when  $R = 61$ . The model clearly underestimates the observed number of trophic links. But what is even worse is that, with  $R = 61$ , the model also underestimates the number of trophic levels. There is a simple explanation for why the base model fails to produce more trophic links and trophic levels. Here we assume that at time  $t$  each species has the same probability of being selected to utilise resources. Imagine at the beginning when all resource are unoccupied (i.e. label value zero), many species will enter the food web to become basal species. This then leads to fewer new species at higher trophic levels resulting in a lower number of trophic levels. Since there are fewer species at higher trophic levels, the number of trophic links naturally will be low. Thus, in the light of the reason given above, we should incorporate some rule such that not all species can be selected to utilise resources with equal probability.

## 2.3. Base model with Matthew effect

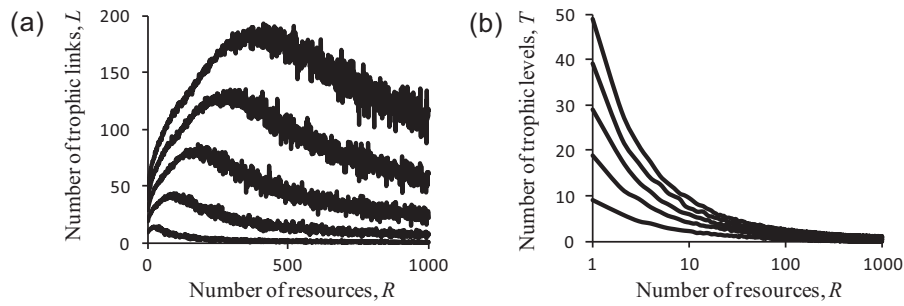
One simple modification of our base model is inspired by network science and sociology. Real networks are dominated by a few hubs that have disproportionately large numbers of connections, and preferential attachment is thought to be a possible mechanism responsible for this phenomenon (Barabási and Albert, 1999). This in essence is the “rich get richer” effect, or better known as the “Matthew Effect” in sociology, as first mentioned by Merton (1968) (after being inspired by the Gospel of Matthew). Similar phenomenon has also been discovered in the analysis of ecological networks where well-connected species (e.g. the generalist predators) tend to be the targets of interaction with new species due to their larger population size or longer active foraging time frames (Olesen et al., 2008; Bascompte and Stouffer, 2009). Thus, after being inspired by all those mentioned above, we incorporate the Matthew effect into our base model as follows. At time  $t$ , we assume the probability of species  $i$  being picked is proportional to the number of resources it has acquired:

$$Pr(i) \propto Q_i^x, \quad (1)$$

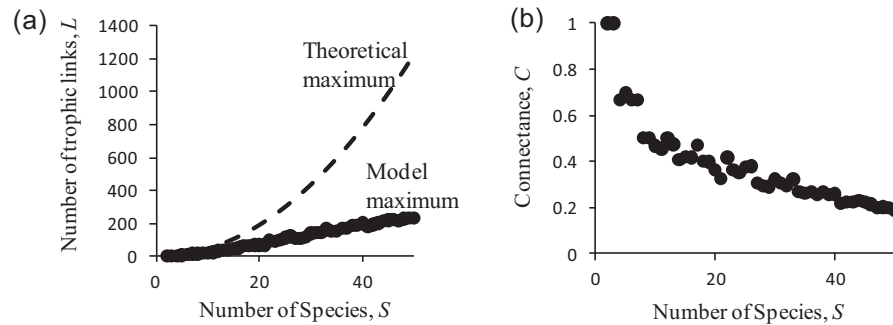
where  $Q_i$  is the number of resources acquired by species  $i$  and  $x$  is a parameter measuring the strength of Matthew effect. When  $x = 0$ , our base model is recovered; and when  $x$  becomes large, species with more resources tend to be picked for utilising other resources.

For a fixed number of species, increasing  $x$  does not result in a greater maximum number of trophic links; but it seems to delay the rate at which the number of trophic links decays to zero when the number of resources approaches infinity (Fig. 6a). Increasing  $x$  slowly results in a higher number of trophic levels, which decays slowly to zero as the number of resources becomes large (Fig. 6b).

As for finding the best parameter values for the Kuosheng Bay food web, we searched through the  $R$ - $x$  parameter space systematically. For each parameter combination, we performed 1000



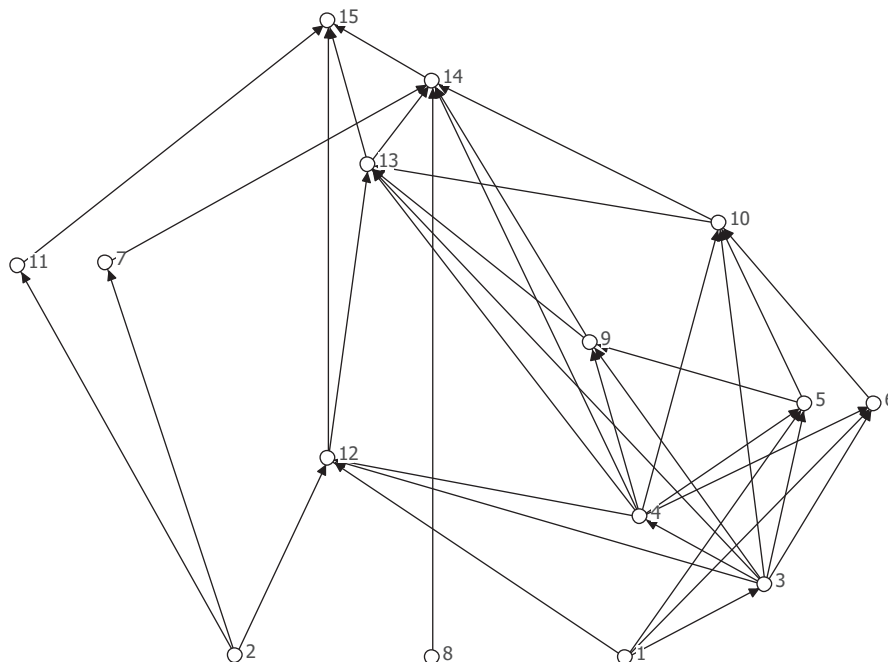
**Fig. 2.** Simulation results for (a) the number of trophic links and (b) the number of trophic levels produced by the base model when the number of resources ( $R$ ) varies from 1 to 1000. In both figures, from the top, the plots are for species number,  $S=50, 40, 30, 20$  and  $10$  respectively. Each data point is the average of 1000 model simulations with a specific set of parameter values.



**Fig. 3.** Simulation results for (a) the maximum number of trophic links attainable and (b) the corresponding connectance as predicted by the base model when the number of species ( $S$ ) varies from 1 to 50. Here, model simulations were carried out with  $R=400$  resources. In (a), the broken link shows the theoretical maximum number of trophic links,  $(S^2 - S)/2$ , given a particular value of  $S$ .

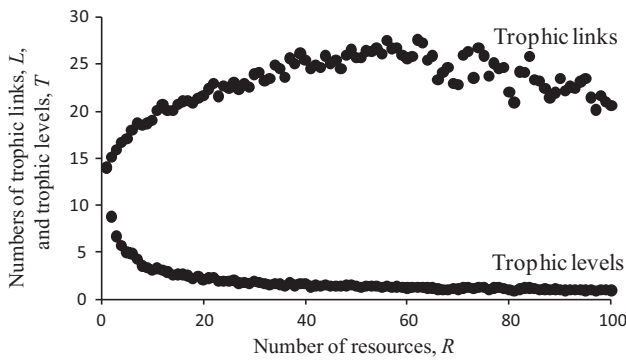
simulations and counted how many times we observed 36 trophic links and 8 trophic levels. We then converted this into the probability of observing the empirical data. Fig. 7 shows the probability of our model producing the observed data in the  $R-x$  parameter space, and the best fit parameter combination here is when  $R=51, x=2$ . We simulated our modified model with the best fit

parameter values 1000 times in order to construct a model distribution for the numbers of trophic links and trophic levels; and the observed data still falls outside the 95% confidence region of the model distribution (Fig. 8a and b). Thus, our base model with Matthew effect cannot reproduce the observed data with meaningful statistical significance.



**Fig. 4.** A graphical representation of the Kuosheng Bay food web in North-East Taiwan. Here nodes are species and directed edges are trophic links connecting resource species to consumer species. In this system, there are 15 species, 36 trophic links and 8 trophic levels.





**Fig. 5.** Numbers of trophic links and trophic levels obtained by simulating the base model with 15 species when the number of resources ( $R$ ) varies from 1 to 100. The maximum number of trophic links produced is 27, and this occurs when  $R=61$ , and corresponds to 2 trophic levels. Each data point is the average of 1000 model simulations with a specific set of parameter values.

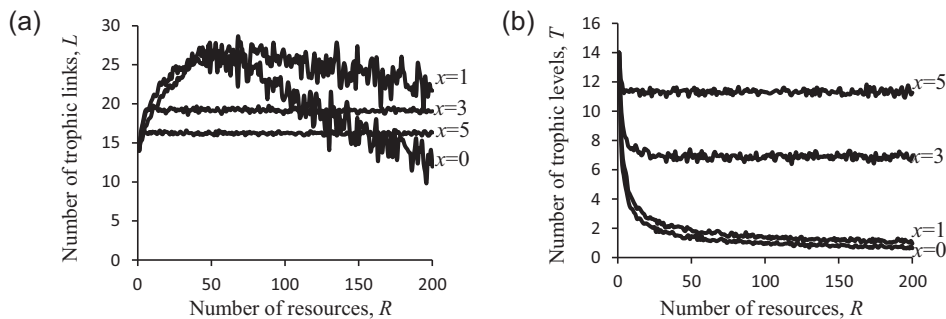
2.4. Base model with utilisation effect

We propose another modification to our base model. In the original model, we assume that resources are equally likely to be utilised by a species if the trophic level of this particular species is higher than the resources' label values. But it has been argued that food quality, in terms of protein and energy required for metabolism and growth of consumers, often increases with the trophic position of resources along the food chain (Bowen et al., 1995; Diehl, 2003). Thus, it might not be far-fetched to assume that resources at higher trophic levels tend to be more preferred than those at lower levels. We modify the base model by incorporating this utilisation effect as follows. At time  $t$ , again we pick a species at random, and then identify what resources are eligible for it to utilise. Next we assume that the probability of an eligible resource  $i$  being picked is proportional to its label value:

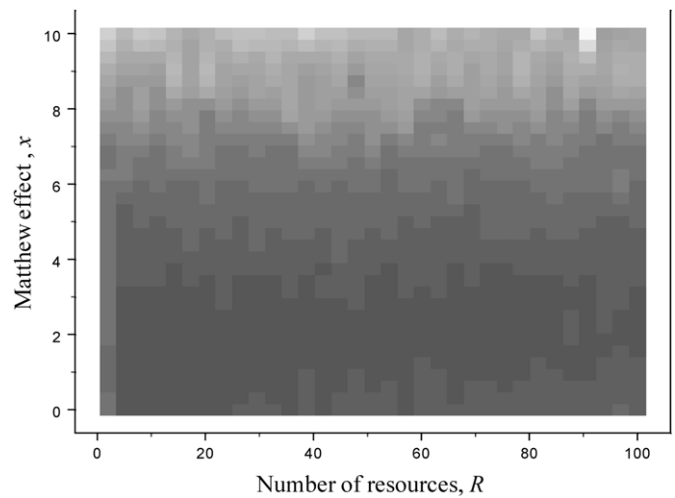
$$Pr(i) \propto V_i^y, \tag{2}$$

where  $V_i$  is the label value of resource  $i$ , and  $y$  is a parameter measuring the strength of utilisation effect. When  $y=0$ , our base model is recovered; and when  $y$  becomes larger, resources of higher label values are more likely to be utilised than those with lower label values.

The effect of increasing  $y$  is to increase the number of trophic links as well as the number of trophic levels (Fig. 9). Unlike the previous modification with Matthew effect, the maximum number of trophic links attainable also increases with larger values of  $y$ ; however, utilisation effect seems not to prolong the decline in the number of trophic links and trophic levels as much as the case with Matthew effect (compare Fig. 9 with Fig. 6).



**Fig. 6.** Results for (a) the number of trophic links and (b) the number of trophic levels produced by simulating the base model with Matthew effect ( $x$ ) when the number of resources ( $R$ ) varies from 1 to 200. In (b), from the top, the plots are  $x=5, 3, 1$  and  $0$  respectively. In both (a) and (b), all simulations were carried out with 15 species, and each data point is the average of 1000 model simulations with a specific set of parameter values.

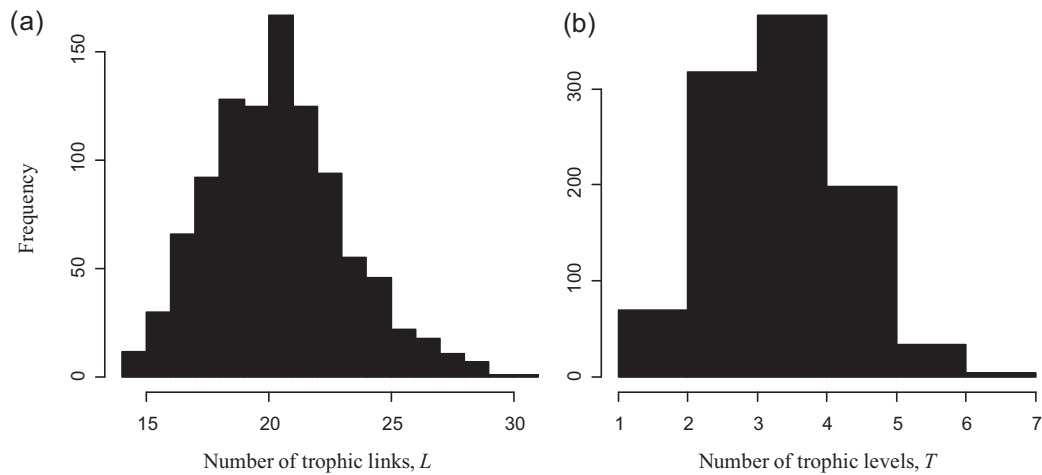


**Fig. 7.** Probability of producing the observed numbers of trophic links and trophic levels as in the Kuosheng Bay food web for various parameter combinations in the  $R$  (number of resources)- $x$  (strength of Matthew effect) parameter space. Grids with darker colour correspond to those parameter combinations with which the model is more likely to produce the observed data. Here, the best parameter values are  $R=51, x=2$ .

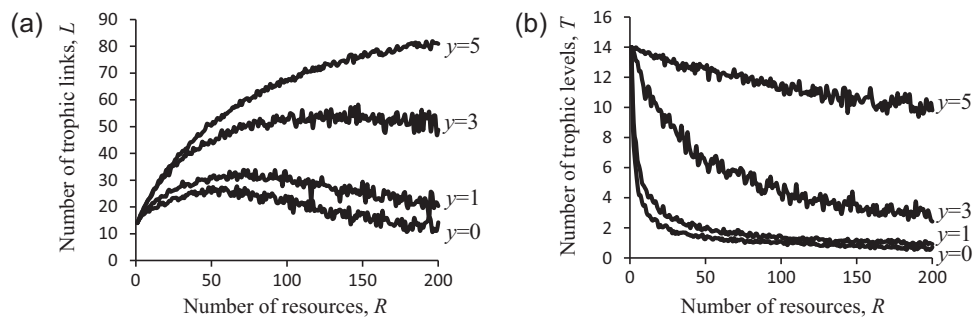
Again, we searched through the  $R$ - $y$  parameter space (Fig. 10), and identified the best parameter values with which the model has the highest probability of producing outcomes similar to the Kuosheng Bay food web (i.e. 36 trophic links and 8 trophic levels). We found the best parameter combination is when  $R=31, y=3$ ; and model simulation with those parameter values shows that the number of trophic links as well as the number of trophic levels observed in the Kuosheng Bay food web are both within the 95% confidence region of the model distribution (Fig. 11).

2.5. Food web structural properties comparison

Next, we take one step further to explore whether this modified food web model (i.e. base model with utilisation effect) can produce other structural properties similar to the Kuosheng Bay food web. We determined the following properties: (1) the number of food chains; (2) mean food chain length; (3) the variation in food chain length; (4) the variation in in-degree (i.e. the number of diet items of a species); (5) out-degree (i.e. number of predators of a species); (6) the proportion of omnivore species (i.e. the fraction of species which have prey from different trophic levels); (7) the number of intra-guild predations (i.e. a structural motif consisting of three species where  $i$  consumes both  $j$  and  $k$ , and  $j$  also consumes  $k$ ); and (8) mean maximum trophic similarity of a food web as defined in Williams and Martinez (2000). We adopted the



**Fig. 8.** Distribution of (a) the number of trophic links and (b) the number of trophic levels obtained when the base model with Matthew effect was simulated 1000 times by using the best set of parameter values (i.e.  $R=51$ ,  $\alpha=2$ ).



**Fig. 9.** Results for (a) the number of trophic links and (b) the number of trophic levels produced by simulating the base model with utilisation effect ( $y$ ) when the number of resources ( $R$ ) varies from 1 to 200. In both figures, from the top, the plots are for  $y=5$ , 3, 1 and 0 respectively. In both (a) and (b), all simulations were carried out with 15 species, and each data point is the average of 1000 model simulations with a specific set of parameter values.

methodology proposed by Williams and Martinez (2000) to examine the fit between the model output and the data: (1) first we simulated the base model with utilisation effect by using the best-fit parameter values 1000 times and then obtained the simulated distribution of a particular food web property; (2) for each food web property, we determined the model mean and its standard deviation; (3) for each food web property, we calculated the difference between the empirical value and the model mean, and then normalised this error by dividing it by the standard deviation of the model distribution; (4) since approximately 95% of simulated values of a food web property are within 2 model standard deviations of the model mean, a normalised error between 2 and  $-2$  thus indicates a good fit between the model and the data (on the other hands, normalised errors greater than 2 or less than  $-2$  are said to be statistically significant and indicate poor fit between the

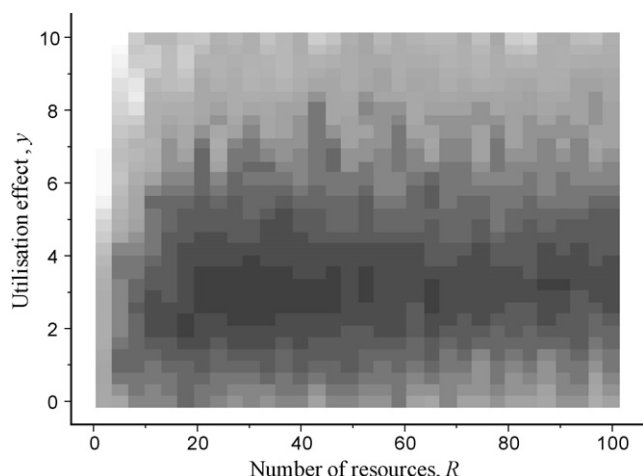
model and the data). To our surprise, the above structural properties derived from the Kuosheng Bay food web can be captured by the model reasonably well as most of the normalised errors are between 2 and  $-2$  (Table 1). However, there is one exception as the model can never produce an out-degree distribution as diverse as the one observed in the real data.

2.6. Further assessment of the base model with utilisation effect

The Kuosheng Bay food web is relatively small in size when compared to other more well-known food webs. In this section, we examine the performance of our base model with utilisation effect in fitting larger food webs and compare the result with those derived from other well-established models.

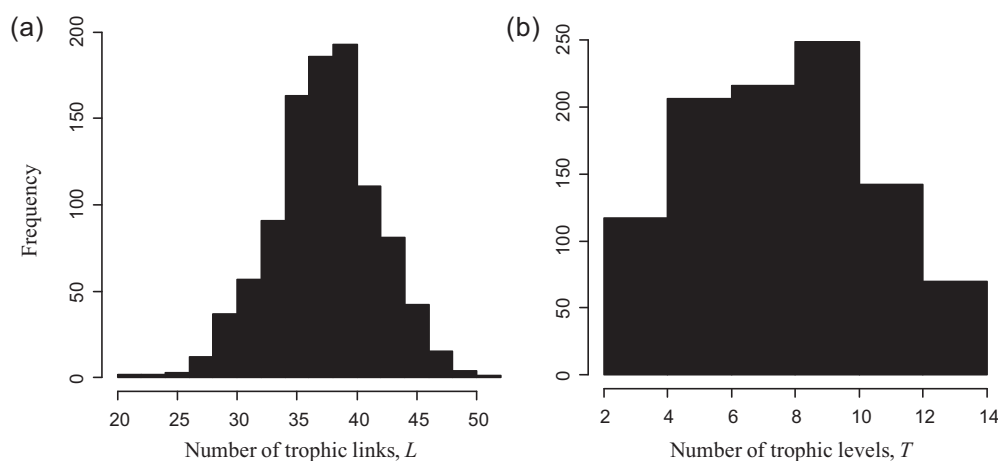
**Table 1**  
 Summary of various food web properties of the Kuosheng Bay food web, as well as their model means, standard deviations (S.D.) and the corresponding normalised errors (N.E.) obtained by fitting the base model with utilisation effect using  $R=31$ ,  $y=3$ .

	Data	Model mean	Model S.D.	N.E.
Number of trophic links	36	37.84	4.09	-0.53
Number of trophic Levels	8	8.07	2.82	-0.15
Number of food chains	59	217	163.38	-1.01
Mean food chain length	4.64	5	1.38	-0.39
Variance of food chain length	1.52	2.17	1.34	-0.59
Variance of in-degree	3.57	5.53	1.22	-1.51
Variance of out-degree	4.11	1.74	0.50	4.71
Proportion of omnivore species	0.53	0.49	0.12	0.19
Number of intraguild predations	28	28.58	9.40	-0.12
Mean trophic similarity	0.50	0.41	0.06	1.69



**Fig. 10.** Probability of producing the observed numbers of trophic links and trophic levels as in the Kuosheng Bay food web for various parameter combinations in the  $R$  (number of resources)– $y$  (strength of utilisation effect) parameter space. Grids with darker colour correspond to those parameter combinations with which the model is more likely to produce the observed data. Here, the best parameter values are  $R=31$ ,  $y=3$ .

We borrowed directly the work from Williams and Martinez (2000) in which they fitted the cascade model and the niche model to seven food web datasets, those are: Skipwith Pond, Little Rock Lake, Bridge Brook Lake, Chesapeake Bay, Ythan Estuary, Coachella Valley and St Martin Island. Twelve food web properties were measured in order to test the prediction power of food web models in that particular study, they are: the proportions of top ( $T$ ), intermediate ( $I$ ) and bottom ( $B$ ) species; the standard deviations of normalised prey counts ( $GenSD$ ) and predator counts ( $VulSD$ ) of species; mean maximum trophic similarity of a food web ( $MaxSim$ ); the mean ( $ChnLg$ ) and standard deviation ( $ChnSD$ ) of food chain lengths, and the log of the number of food chains ( $ChnNo$ ) (a food chain here is defined as a path from a species to a basal species, and in this section we adopt this definition for the ease of comparison and consistency); and finally the proportion of omnivore species ( $Omniv$ ). We have to note that Williams and Martinez (2000) also quantified the proportion of cannibal species and the proportion of species involved in longer feedback loops in their study; but these two properties were omitted in our study here due to the fact that simulated food webs from our model are all acyclic digraphs (i.e. graphs with no self-loops and feedback loops).



**Fig. 11.** Distribution of (a) the number of trophic links and (b) the number of trophic levels obtained when the base model with utilisation effect was simulated 1000 times with the best set of parameter values (i.e.  $R=31$ ,  $x=3$ ).

For each of those seven food webs, we searched systematically through the  $R$ – $y$  parameter space and identified the best parameter values with which our model is most likely to produce the observed number of trophic links and trophic levels (best parameter values are summarised in Table 2). We then simulated the base model with utilisation effect by using the best-fit parameter values 1000 times; and for each of the above mentioned food web properties, we again calculated its normalised error to assess the performance of our model (Table 3).

In total, there are 66 normalised errors, and their mean and standard deviation are 2.04 and 5.59 respectively. As derived from Williams and Martinez (2000), the mean and standard deviation of the corresponding normalised errors for the cascade model are  $-3.04$  and  $14.10$  respectively while their counterparts for the niche model are  $0.11$  and  $1.91$  respectively. Furthermore, 22 out of those 66 normalised errors for our model are greater than 2 standard deviations of their model means, and the corresponding numbers for the cascade model and the niche model are 47 and 15 respectively. Thus, all these results suggest that the performance of our model is better than the cascade model but worse than the niche model. A close inspection of Table 3 suggests that our model performs badly in predicting the proportion of top species ( $T$ ) and the variation in normalised predator counts ( $VulSD$ ).

### 3. Discussion

In this paper, we have constructed a model for food web by adopting a very simple rule: a species encounters and utilises a suitable resource by random chance, and this species then subsequently becomes a resource for other species where trophic interaction then occurs. The assembly process of a simulated food web roughly goes as follows. At the beginning all resources are unoccupied, and as new species enter the system, they deprive those unoccupied resources to become basal species. The number of unoccupied resources then declines while the number of resources at first trophic level increases (in the disguise of those basal species). As a consequence of this, future new species are more likely to encounter occupied resources, and resources at higher trophic levels (in the disguise of those newly introduced species) start to bloom. In general, the food web grows as new species gradually enter the system and pile on top of existing species forming all sorts of trophic interactions. A natural consequence of this is that a new species is less likely to occupy resources or consume species that have trophic levels much lower than its own; and its

**Table 2**

Summary of estimated parameter values after fitting the base model with utilisation effect to various food web datasets in Williams and Martinez (2000).

Parameter values	Skipwith Pond	Little Rock Lake	Bridge Brook Lake	Chesapeake Bay	Ythan Estuary	Coachella Valley	St Martin Island
R	440	975	120	40	225	500	180
y	4	2.9	3	1	2	4	2.5

diet range should constitute those species occupying trophic levels immediate below (or a few levels below).

The base model takes on two parameters, the number of species and the number of resources. Unlike other food web models in the literature, our model does not require species being ordered hierarchically in some dimensions a priori, nor does it require a pre-defined number of trophic links. For a fixed number of species, our model can predict the number of trophic links and its maximum limit; and it further demonstrates that there is a natural negative relationship between the number of species and food web connectance. This finding may also shed some lights on an old debate on whether food web connectance increases with the number of species. Traditionally, both from theoretical and empirical grounds, food web connectance should have a negative relationship with the number of species. A theoretical study by May (1972) demonstrates that an ecosystem is only stable if the product of species number and connectance is less than the inverse of the average interaction strength between species pairs; thus, as species number increases, connectance should decline in order to fulfil the stability condition of May's theoretical model (Begon et al., 2006). Briand (1983) compiled several food web data from the literature and found that connectance decreases non-linearly with the number of species. Our base model here does demonstrate the negative and non-linear relationship between connectance and species number in a manner similar to Briand's (1983) empirical findings. However, in contrast to May's theoretical work, our analysis here does not have to go into the complicated issue of system stability in order to demonstrate this relationship. Naturally and simply, as our base model has demonstrated, the decline in connectance with increasing species number could simply be the end result of chance encounter between species and resources that sets a limit on the

amount of trophic interactions observable. Having said all this, there are also studies demonstrating that the relationship between connectance and species number is anything but negative (Warren, 1989; Winemiller, 1990; but see Gilbert, 2009; Poulin, 2010). One hypothesis for this is based on the behavioural and morphological characteristics of species (Begon et al., 2006). The hypothesis assumes that each species has its own diet range and is able to consume all species that fall within this range. Naturally, if species richness increases, then the number of species within a particular species' diet range will also increase; and this will result in more trophic interactions and thus will keep connectance constant at least. In contrast, if a species only feeds on a fixed number of species, then connectance is bound to fall as species richness increases. At present, each species in our model does not have a diet range; and this might be a reason for why our model cannot account for the non-negative relationship between connectance and the number of species, and therefore limits the generality of our model.

One other finding from our base model is that the number of trophic levels always declines with ever-increasing number of resources. Empirical studies have shown that food chains tend to shorten along the gradient of resource enrichment after reaching the maximum at intermediate resource availability (Diehl and Feissel, 2001). A possible simple mechanism is that intra-guild predators would encounter and consume proportionally more basal resources than their intra-guild preys when the amount of available basal resources increases (Post, 2002; Post and Takimoto, 2007). Although we did not explicitly monitor the change of food chain length, which is determined by the trophic position of the top predators in our model, but similar mechanism may apply to our simulation results here as well. What our model fails to capture is the positive relationship between food chain length and resource

**Table 3**

Model means for various food web properties after the base model with utilisation effect was fitted to each of those food web datasets in Williams and Martinez (2000). Each model mean is followed by two numbers in parenthesis: first is the standard deviation of the model distribution and the second is the normalised error. Normalised errors greater than 2 or smaller than -2 are in bold, indicating poor fit between the model and the data. NA stands for not applicable as the empirical properties were not given in Williams and Martinez (2000).

Food web Property	Skipwith Pond	Little Rock Lake	Bridge Brook Lake	Chesapeake Bay	Ythan Estuary	Coachella Valley	St Martin Island
T	0.05 (0.02, -0.36)	0.02 (0.01, -0.50)	0.05 (0.02, -2.24)	0.05 (0.03, <b>8.70</b> )	0.02 (0.01, <b>32.5</b> )	0.04 (0.01, -3.10)	0.03 (0.02, <b>7.14</b> )
I	0.47 (0.18, <b>2.57</b> )	0.42 (0.09, <b>4.77</b> )	0.57 (0.12, 0.92)	0.47 (0.07, 0.73)	0.50 (0.06, 0.73)	0.55 (0.15, <b>2.31</b> )	0.49 (0.11, 1.88)
B	0.48 (0.50, -0.88)	0.57 (0.50, -0.88)	0.38 (0.49, -0.12)	0.48 (0.50, -0.63)	0.48 (0.50, -0.78)	0.41 (0.49, -0.63)	0.48 (0.50, -0.67)
GenSD	1.12 (0.31, -0.65)	1.35 (0.17, 0.40)	1.04 (0.16, 0.30)	1.29 (0.14, -3.60)	1.35 (0.09, -2.30)	1.00 (0.23, -1.18)	1.23 (0.16, -1.29)
VulSD	0.46 (0.04, 1.83)	0.40 (0.03, <b>6.87</b> )	0.46 (0.05, <b>3.21</b> )	0.55 (0.07, <b>8.63</b> )	0.53 (0.04, <b>23.52</b> )	0.45 (0.03, <b>5.55</b> )	0.46 (0.04, <b>8.87</b> )
MaxSim	0.81 (0.09, -0.55)	0.42 (0.03, <b>10.29</b> )	0.48 (0.05, <b>5.18</b> )	0.42 (0.06, 1.31)	0.34 (0.02, <b>7.89</b> )	0.73 (0.08, -0.12)	0.42 (0.04, <b>2.77</b> )
ChnLg	5.05 (2.58, 0.45)	7.89 (4.32, NA)	5.39 (2.05, -0.66)	2.98 (0.70, 1.44)	5.97 (2.09, -0.03)	6.88 (2.76, -0.07)	5.38 (2.06, -0.09)
ChnSD	1.38 (0.52, 0.10)	1.61 (1.25, NA)	1.50 (0.50, -1.14)	0.93 (0.20, 1.35)	1.27 (0.40, 0.47)	1.69 (0.52, -0.46)	1.34 (0.44, -0.08)
ChnNo	4.21 (1.20, -0.41)	7.18 (2.07, NA)	3.85 (0.72, -1.39)	2.55 (0.29, -0.60)	5.14 (0.88, -1.26)	5.24 (1.24, -0.75)	4.42 (0.84, -1.07)
Omniv	0.40 (0.21, 0.96)	0.29 (0.12, NA)	0.47 (0.15, -0.46)	0.24 (0.08, <b>3.47</b> )	0.33 (0.07, <b>2.98</b> )	0.48 (0.19, 1.49)	0.36 (0.13, 1.91)



number at very low resource availability (Post, 2002). A possible explanation for this discrepancy is that we did not consider the loss of resources that comes along with trophic interaction in the food web assembling process. Further refinements of the model should incorporate an additional parameter for the transfer efficiency of resources across different trophic levels (Pauly and Christensen, 1995), and one should investigate how this can improve our model.

In this paper, we also confronted our base model with a small empirical food web and found that the model underestimated the number of trophic links and the number of trophic levels. We proposed two modifications to the base model that have some ecological/biological relevance to nature, and they might shed some lights on how real ecosystems (at least for our small Kuosheng Bay food web) might grow over time. First modification assumes that species with more resources are likely to acquire more resources (i.e. Matthew effect). With this mechanism, resources at one trophic level will be taken up quickly by a few species; and this results in an increase in the number of trophic levels (as new species will have to look elsewhere but lower trophic levels for food), but not necessarily results in a higher number of trophic links. However, our base model with Matthew effect fits the real data poorly, and this suggests that the assembly process of the Kuosheng Bay ecosystem is unlikely to follow this route. Next, we then proposed a second modification to our base model where resources at higher trophic levels are more likely to be taken up than lower ones (i.e. utilisation effect). This mechanism dictates the assembly process of a food web in a manner such that resources at one trophic level do not have to be deprived completely before the appearance of species at higher trophic levels. In the base model, the majority of basal species have to appear before the emergence of herbivores, so as the majority of herbivores have to appear before the appearance of carnivores and so on. But with this second modification, species at higher trophic levels can appear even before lower trophic levels are being populated by lower species. With this mechanism in place, our model can now fit the real data better and this suggests that the Kuosheng Bay food web might be assembled in this way.

Due to the small size of the Kuosheng Bay food web, the fit between our base model with utilisation effect and the empirical data might be simply due to chance alone. Thus, we further assessed the performance of our model by fitting it to seven much larger food webs and compared the result with those derived from other food web models. Although our model can fit those empirical data reasonably well, but it falls short in predicting the proportion of top predators. We suspect that this discrepancy between the model and the data may be due to the stopping rule of our model. Remember that our model requires all resources being occupied by species at the highest trophic level in the end of food web simulation. Under such a rule, all top predators are those species at the highest trophic level and this might be an unwanted artefact that renders the model predicting poorly the proportion of top predators. One way to remedy this, and at the same time making our model more general, is to incorporate the loss of resources into future models. With such a mechanism in place, not all resources will be passed on to those species at the highest trophic level. This certainly will produce several top predators at various trophic levels and thus will make the proportion of top predators more variable. Another shortcoming of our model is that the model out-degree variation is consistently lower than those observed in the data. This implies that prey species in our model have more equal chance of being consumed than those observed in the empirical data. One way of introducing more heterogeneity in the risk of predation is by combining Matthew effect and utilisation effect in the same modelling framework. With Matthew effect in place, some species will have more resources than the majority others; and this will render them being consumed by many species at higher trophic levels resulting in an increase in out-degree variation. Having said this, we

are yet to come up with a natural way of simultaneously incorporating both effects in our base model, and this will be an open problem for future investigation. In comparison with other food web models, we found that our base model with utilisation effect performs better than the cascade model, but it still lags behind the niche model. Basing on the mean normalised error, the fit between data and the niche model is considerably better than ours. We have to note that two out of 66 normalised errors from our model are extremely large and omitting them results in a much smaller mean normalised error. If we consider only the number of significant normalised errors (i.e. those greater than 2 or less than -2), then the gulf between the performances of our model and the niche model becomes smaller (22 out of 66 compared to 15 out of 66).

To sum up, in this paper we have proposed a fundamental framework for food web modelling that takes a different approach from the mainstream led by cascade-related models. We believe this model is of a more fundamental nature than the existing ones as it regards the hierarchical ordering of species and the number of trophic links as two outcomes generated by the fundamental processes embedded in the model. However, our work is far from complete. Surely, the next logical step is to modify our base model in order to reflect more realistic biology, and the fit between our model and the data can only get better. On the other hand, we believe an equally challenging research direction here is not to make the model more realistic, but is to look for a model which is even more fundamental than what we have proposed here. Finally, last but not the least, we feel the process assumed in our model resembles how one sets up a fish tank from scratch. As most aquarium keepers are aware, to build a moderately self-sustained fish tank, one should start with culturing microbes and algae in the water; this is then followed by adding some algae feeding crustaceans or snails before putting in any carnivorous fish in the tank. And to this end, we thus have justified the name of our model as shown in the title of this paper.

## Acknowledgements

We are grateful for the Organisers of ECEM 2011 in Riva del Garda, Italy, for the possibility to present our work in the conference. We also thank two anonymous reviewers for their constructive suggestions and criticisms.

## References

- Allesina, S., Alonso, D., Pascual, M., 2008. A general model for food web structure. *Science* 320, 658–661.
- Bascompte, J., Stouffer, D.B., 2009. The assembly and disassembly of ecological networks. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1781–1787.
- Barabási, A.L., Albert, R., 1999. Emergence of scaling in random networks. *Science* 286, 509–512.
- Begon, M., Townsend, C.R., Harper, J.L., 2006. *Ecology: From Individuals to Ecosystems*, fourth ed. Blackwell Publishing.
- Bowen, S.H., Lutz, E.V., Ahlgren, M.O., 1995. Dietary protein and energy as determinants of food quality: trophic strategies compared. *Ecology* 76, 899–907.
- Briand, F., 1983. Environmental control of food web structure. *Ecology* 64, 253–263.
- Cattin, M.F., Bersier, L.F., Banasek-Richter, C., Baltensperger, R., Gabriel, J.P., 2004. Phylogenetic constraints and adaptation explain foodweb structure. *Nature* 427, 835–839.
- Cohen, J.E., Briand, F., Newman, C.M., 1990. *Community Food Webs: Data and Theory*. Springer-Verlag, Berlin, New York.
- Cohen, J.E., Pimm, S.L., Yodzis, P., Saldana, J., 1993. Body sizes of animal predators and animal prey in food webs. *Journal of Animal Ecology* 62, 67–78.
- Diehl, S., 2003. The evolution and maintenance of omnivory: Dynamic constraints and the role of food quality. *Ecology* 84, 2557–2567.
- Diehl, S., Feissel, M., 2001. Intraguild prey suffer from enrichment of their resources. A microcosm experiment with ciliates. *Ecology* 82, 2977–2983.
- Dunne, J.A., Williams, R.J., Martinez, N.D., 2002. Food-web structure and network theory: the role of connectance and size. *Proceedings of the National Academy of Sciences of the United States of America* 99, 12917–12922.

- Dunne, J.A., Williams, R.J., 2009. Cascading extinctions and community collapse in model food webs. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1711–1723.
- Estrada, E., 2007. Food webs robustness to biodiversity loss: the roles of connectance, expansibility and degree distribution. *Journal of Theoretical Biology* 244, 296–307.
- Gilbert, A.J., 2009. Connectance indicates the robustness of food webs when subjected to species loss. *Ecological Indicators* 9, 72–80.
- Jordán, F., Liu, W.-C., Mike, A., 2009. Trophic field overlap: a new approach to quantify keystone species. *Ecological Modelling* 220, 2899–2907.
- Lin, H.J., Shao, K.T., Hwang, J.S., Lo, W.T., Cheng, I.J., Lee, L.H., 2004. A trophic model for Kuosheng Bay in Northern Taiwan. *Journal of Marine Science and Technology* 12, 424–432.
- May, R.M., 1972. Will a large complex system be stable? *Nature* 238, 413–414.
- Melián, C.J., Bascompte, J., 2002. Complex networks: two ways to be robust? *Ecology Letters* 5, 705–708.
- Merton, R.K., 1968. The Matthew effect in science. *Science* 159, 56–63.
- Olesen, J.M., Bascompte, J., Elberling, H., Jordano, P., 2008. Temporal dynamics in a pollination network. *Ecology* 89, 1573–1582.
- Pauly, D., Christensen, V., 1995. Primary production required to sustain global fisheries. *Nature* 374, 255–257.
- Pimm, S.L., Lawton, J.H., Cohen, J.E., 1991. Food web patterns and their consequences. *Nature* 350, 669–674.
- Post, D.M., 2002. The long and short of food-chain length. *Trends in Ecology & Evolution* 17, 269–277.
- Post, D.M., Takimoto, G., 2007. Proximate structural mechanisms for variation in food-chain length. *Oikos* 116, 775–782.
- Poulin, R., 2010. Network analysis shining light on parasite ecology and diversity. *Trends in Parasitology* 26, 492–498.
- Rossberg, A.G., Matsuda, H., Amemiya, T., Itoh, K., 2006. Food webs: experts consuming families of experts. *Journal of Theoretical Biology* 241, 552–563.
- Warren, C.P., Pascual, M., Lafferty, K.D., Kuris, A.M., 2010. The inverse niche model for food webs with parasites. *Theoretical Ecology* 3, 285–294.
- Warren, P.H., 1989. Spatial and temporal variation in the structure of a freshwater food web. *Oikos* 55, 299–311.
- Williams, R.J., Martinez, N.D., 2000. Simple rules yield complex food webs. *Nature* 404, 180–183.
- Winemiller, K.O., 1990. Spatial and temporal variation in tropical fish trophic networks. *Ecological Monographs* 60, 331–367.
- Woodward, G., Ebenman, B., Emmerson, M., Montoya, J.M., Olesen, J.M., Valido, A., Warren, P.H., 2005. Body size in ecological networks. *Trends in Ecology & Evolution* 20, 402–409.