Evaluation of the community response of ecological networks using complexity science

by

Xueke Lu

Submitted for the degree of Doctor of Philosophy

School of Electronic Engineering and Computer Science
Queen Mary, University of London
United Kingdom

June 2016

Statement of originality

I, Xueke Lu, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged at the end of this thesis.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature: Xueke Lu

Date: 15/06/2016

Abstract

This thesis investigates network properties of natural food webs. In particular, it focuses on the effect that external disturbances have on their substructures and robustness. The importance of a network-level methodology lies in its capacity to capture entangling species interactions and identify inter-connecting properties in heterogeneous food webs.

The research first analysed the responses of freshwater food webs under the stress of drought. A core/periphery structure was detected and its relative size was found to be unchanged after drought despite a significant biodiversity loss. Species extinction triggered extensive link rewiring and movement of species from the core to the periphery. These results showed that the robustness was maintained indicating that the redundancy in the core can effectively mitigate species level perturbations. Secondly, the research further examined the effects of Genetically Modified Herbicide Tolerant (GMHT) management on food web properties and robustness. Network analysis showed that such change in farming practice has no significant impact on the agro-ecosystems. However, crop switching, a common practice in agriculture, was found to pose much more significant changes on network properties and robustness when compared to GMHT crops. Thirdly, the research examined over 50 empirical food webs and demonstrated that the relative core size is a much more effective indicator of food web robustness than the classical ecological measure connectance, as the latter was found to be insensitive to changes in the interaction patterns. Lastly, the research established the relationships between centrality measures and species ecological and/or functional role in food webs, and how they impact on network robustness.

TO MY FAMILY

Acknowledgments

First of all, I thank Athen Ma and Steve Uhlig who gave me an offer of this PhD position. During which I have been given invaluable guidance from my main supervisor Athen, who have taught me how to conduct research and write academic papers and manage multiple tasks and handle problems. I also thank my co-supervisors Steve, Guy Woodward and Raúl Mondragón, who always support me and and provide pertinent guidance on research whenever I need them. I thank Chinese Scholarship Council (CSC) and Queen Mary who have funded my PhD.

Additionally, I was given the opportunities to collaborate with Mark E ledger, David A Bohan, and other brilliant people. I also learned quite a lot through meetings and discussions with Eoin O'Gorman, Lei Zhao, and Samuel Johnson. Those experiences were not always available for a PhD student, and I really appreciate it.

Finally, I would like to thank my colleague Clare Gray, who has helped me countless times on my English writing and my Thesis. I also thank Weisi Guo, Haroon Mumtaz, Hamed Haddadi, Sam, Raúl, and Steve, who helped me with my thesis and the proofreading. In addition, I thank my friends Shenglan Huang, Anqi He, Dantong Liu, Xiang Xu, Xinyue Wang, Hanbing Mu, Yunong Pang, Fei Dong, Liran Sun, Yanxu Meng, and Momo Zheng, who have helped me in many ways, and looked after me whenever I need them. Last but not least, I thank my parents in China, who have unconditionally supported me during the past four years.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Food webs

1.1.1 The emergence of a network approach

Complex networks in ecology contain species (nodes) that exist within an ecosystem and the interactions among them (links). Ecological networks recently have been highlighted as a useful tool for various purposes, including bio-monitoring [1] (Lu, 2014), conservation [2–5], and theoretical ecology advances [6, 7]. A network approach is important for ecology: First, a network approach can help explain the switch of ecological status of ecosystems [1, 8]. For example, external perturbations can switch lake ecosystems from pelagic-dominated (i.e., consists of only those species that distributed in the open water area) to benthic-dominated (i.e., consists of only those species that distributed at the lowest level of freshwater systems) [9, 10]. This can be explained using energy flow through trophic interactions, indicating that regardless of the incoming resource types of the ecosystems, the consumers can survive as long as the interaction links are preserved [1, 8, 11]. Second, by analysing multiple food webs from different ecosystems, the stability of food webs has been found to depend more on the 'weak interactions'

[12–14]. Other features include the species interaction pattern [15] and the important rewiring ability [16], both of which can be further correlated with how robust a food web is. Finally, a network approach can help identify key species that are more important than others in conserving the ecosystem stability and the robustness under perturbation through the analysis of species interactions [17–21]. For example, the generalist predator (i.e., with many prey) has a controlling effect on the food web structure, and the removal of which will cause the ecosystem to break down [22]. Similarly, a prey species with a large biomass can also be seen as a keystone species, the removal of which can cause cascade effects on most of the consumers who lost their prey [23].

1.1.2 Existing network analysis on food webs

Community is an important concept in both network science and ecology, and it can have very different meanings. In network science, a community usually represents a subset of the network which is either densely connected or functionally related, or both [24–28]. In ecology, a community is often referred to a group of species that are geographically located in the same area during a specific time and therefore might have direct or indirect interactions with each other [13, 18, 29, 30].

Finding suitable ways to measure and predict ecosystem stability and community responses is a fundamental challenge in ecology. In order to study the ecosystem and its functioning, a common way is to evaluate the properties of species that exist within the ecosystem. Traditional food web properties include individual species biomass change [31, 32], species interaction strength [33, 34], and the effect of keystone species removal [3, 18]. For example, drought can significantly reduce the biomass production of relatively large sized species, whilst the species that are small are able to adapt quickly [35]. It has also been observed that not all of the large predators went extinct under stressors [36], therefore, there must exist other factors that affect a species survivability apart from individual species properties. The species level approaches to study ecosystem responses to environmental stressors can only demonstrate the consequences, while the cascade effects or the stabilising factor from the other species cannot be detected only through the analysis of individual species properties.

Network properties of a food web can be grouped into two aspects. First, the properties that associate with individual species, including the degree centrality, which represents its importance within an ecosystem [17, 37]; the generality, which describes how many prey it has; and the vulnerability, which describes how many predators that depend upon it [5, 38, 39]. Second, the network-level properties such as the connectance, describes the link density [40–43]. The proportion of basal / intermediate / top-level species indicates the balance among the number of resource, the intermediate energy transmit hubs, and the top predators [38, 44], where basal species (those with only outgoing links) form the bottom layer of the network, top-level species (those with only incoming links) form the top layer of the network, and the rest of the network consist of intermediate level species (Figure 1.1).

Food webs in general are well connected, and display a hierarchical structure known as a trophic level structure [12, 16, 18–20]. Trophic level describes the position of the species in a food chain and their predatory behaviours [45, 46]. The trophic height of a food web is also related to the efficiency of energy flow. A higher food chain lengths generally indicates higher energy availability and productivity within systems [47, 48], it also indicates the food web complexity, therefore has been used as an indicator of species predation behaviour change under different environment conditions [36, 49]. For example, drought caused top predators to go extinction which reduced the trophic height of the food webs [36]. Another network property is a recently developed metric called trophic coherence, which correlates food web stability with restrained species trophic position diversity of the whole network, indicating that certain level of similarity in the species trophic position can provide flexibility when perturbation happens [46].

Food webs are shown to have non-random network structures from various perspectives [12, 15, 16, 18, 19, 19, 20], however, few studies explored the response of food web under stress through the changes in their structures or species connectivity patterns.

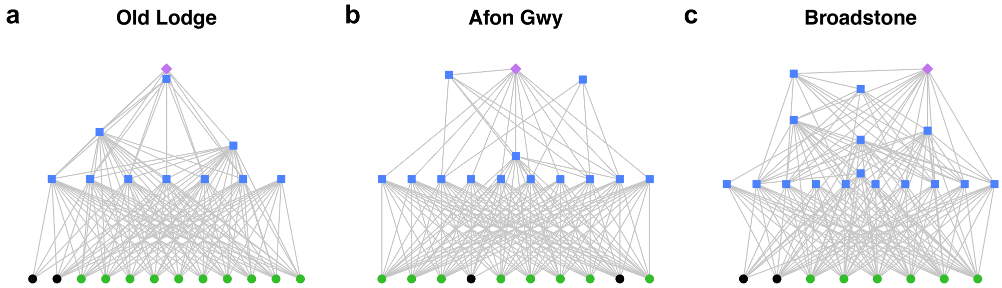


Figure 1.1: Three freshwater food webs (**a** - Old Lodge, **b** - Afon Gwy, **c** - Broadstone) plotted in trophic levels. Black nodes are detritus, green nodes are producers, and the rest are higher level predators. Predators, especially those from the second trophic level, share similar connectivity patterns with their prey.

For example, recent studies on how climate change (i.e., drought) would affect food web topology and functional properties captured significant biodiversity change and biomass flux change [35, 36, 50]. However, the methods used in these studies cannot explain why connectance of the food web was conserved while food web size shrunk significantly. Also, questions such as what would be the species rewiring pattern under environment change still remain to be answered. These examples show that there exist limited network approaches to fully understand the community responses under environment perturbations. It is important to cross-fertilise ecology and network science by utilising advanced network metrics to help better understand food web structure and properties.

1.1.3 Network analysis on substructures

There have been an increasing interest on using substructure network analysing metrics to examine a network [24, 25, 27, 51–53]. A large and densely connected core allows certain level of flexibility and adaptability when perturbations occur [54], whilst a smaller core helps with controllability of the whole network [55, 56]. The core/periphery structure has been detected in many types of networks, such as the underground networks [57], the brain networks [58], and social networks [59]. In all the cases, the core is stable through time and provides efficiency in communication with all the other nodes in the network.

The reasons that network substructure analysis on food webs has become a relatively uninvestigated area are twofold. First, food webs are much smaller compared with other kinds of complex networks [60], such as human brain networks (up to billions of nodes) and the World-wide-web (around billions of hyperlinks). Therefore, it is not always possible to detect community structure using existing metrics that were more suitable for larger networks [6, 61–64]. Food webs consist of species that can be grouped based on different trophic levels, and the same level species always form very similar connectivity patterns (e.g., Figure 1.1), which effectively makes most of the food webs un-dividable into modules. Second, food web substructure analysis focuses on functionality of nodes and the interaction strength of links which neglect the actual connectivity pattern [14, 15, 51]. For example, compartments can be detected in the Chesapeake Bay food web that consist of different functional communities, such as the pelagic compartment (in the water area) and the benthic compartment (in sediments) [51]. However, this compartment method is only effective if the food webs contain different functional communities [51]. Yet, community structures that explain the underlying mechanism that buffers perturbations under environment stressors still remains to be identified.

1.2 Food web robustness

Attacks which involve the targeted removal of nodes are able to cause catastrophic effects. This is inspired by the *percolation theory*, which describes the effect of node removal is able to percolate through the network, with direct effect being the consequential removal of all the links associated with the removed node [62, 65–70]. The strategy to trigger node or link removal varies. In this context, a node removal is a so called *site percolation*, whereas a link removal corresponds to a *bond percolation* [67].

Consider a malicious attack on the most important node with the highest degree (e.g., the central server, which connects to the largest number of other computers in the Internet infrastructure [71]), the effect of attacks is able to propagate quickly from this highest degree node to all its neighbours. That same consequence would apply to an attack caused by a virus, which will spread much faster than if a random node were to be infected [67, 72, 73].

A key assumption when modelling food web robustness is to define the criteria for secondary extinction. A unified assumption for all the varieties of robustness metrics is the *bottom-up* effect. In ecology, *bottom-up* effect represents the status that bottom-level resource can have an influence towards all the species from a higher level through the connected links among them [5, 17, 74, 75], and it is defined as: if a predator loses all its prey [17], then the predator goes extinct. *Top-down* effect reflects the pressure from top predators on its prey survivability and eventually affects other predators that rely on those lower level prey [5, 75, 76]. The top-down effect has found its application in many population dynamics models, which consider various kinds of propagation effects on population change [5, 75, 76].

Connectance as a network metric, is often used to obtain high level understanding of the network composition. It represents the probability that a link exists between any two selected species [38, 42, 77]. Connectance (C) ranges from 0.06 to 0.31 for most food webs [6, 15, 42]. Within the range, webs with $C < 0.07$ are considered as extremely low connectance webs, which are in general fragile to species extinctions [42]. While $0.07 < C < 0.14$ are intermediate values of connectance, and food webs within this range are much more robust to species extinctions [17, 42]. $C > 0.15$ is the boundary for high connectance food webs, whose robustness to species extinction is among the highest [42]. Theoretical and empirical studies showed that increasing the connectance led to an increase in food web robustness, both from population dynamics view [78, 79] and topological view [4, 16, 17, 42, 80].

Yet, connectance alone has limited power in detecting various environmental perturbations on food webs. When a perturbation presents itself upon the network, connectance fails to indicate any substructural changes that might occur as a result. For example, food webs can respond to the stressors (i.e., pH [81] and drought [36, 50]) by re-generating or establishing new links, yielding a similar level of connectance, whilst changing its topology.

1.3 Node centrality metrics

Node centrality is an important aspect of network analysis, which aims to identify the most important nodes. Many centrality measures exist because different applications may find one set more suitable and logical than another. One major application area of node centrality in food webs is to simulate species removal based on centrality ranking and predict robustness [16, 17, 80, 82–84]. In ecology, a high degree species is important to the whole food web since it governs the most energy flow through the network [20, 85, 86]. The removal of high degree species results in removing a large number of links, which could cause a significant effect on the network. Apart from degree-based removal, other types of centrality based removal exist. For example, recent studies evaluated the consequences of targeted removal based on different centrality metrics on a set of empirical food webs [17, 19, 80, 87]. The focus mainly is to compare and choose the centrality metric that can break down the food web in the most effective way. Studies that focus on how centrality can be used to characterise nodes with different ecological or functional meanings remain scarce. For example, the node with the highest betweenness in protein-protein interaction networks acts as an important agent between modules and this node normally takes part in multiple tasks [88, 89]. But what are the common properties of high betweenness nodes in food webs? Do they share similar properties?

1.4 Research novelty

Firstly, a core/periphery structure has been identified in a set of freshwater stream mesocosms food webs, and it has been shown that the core governs the energy flow against perturbations. Drought reduced the link density within the core whilst the relative size of the core was preserved. The preservation of the core size has buffered the effect of drought and hence food web robustness is unaffected.

Secondly, a comprehensive network analysis has been applied on agroecosystem for the first time to assess the impact of GMHT management. The results presented in this thesis have shown that GMHT management does not significantly affect species interactions in food webs, while switching crop, a common practice in farming, has much more profound effect on network properties and robustness.

Thirdly, the relative core size has been shown as a more effective indicator of food web robustness than connectance, as the core is key in providing food web redundancy which was quantified by constructing spanning trees. The core has also been found to be a fully functioned substructure, as the removal of all the peripheral nodes has little impact on the network integrity.

Lastly, the effectiveness of centrality measures in characterising nodes with different ecological properties was examined. It has been demonstrated that the degree, eigenvector and closeness centrality are linked to species with specific ecological properties, which were previously unknown.

1.5 Thesis organisation

Chapter 2 consists of a detailed description of the methods that are relevant to or used in this research. It includes food web analysis, network substructure analysis, food web robustness measurements, statistical tools, and centrality metrics.

Chapter 3 explores why food webs can preserve their network structure under stress by analysing the substructure properties. A rich core structure is observed in all food webs from both normal and stressed ecosystems, the core link density, species movement and extinction from either core or periphery, and the food web robustness are examined [90] (Lu, 2016).

Chapter 4 compares the community responses under two different stressors: crop switching and GMHT management practices by analysing a set of agricultural food webs. The responses are gauged from species level, substructure level, and the whole network level respectively.

Chapter 5 utilises 53 food webs from different ecosystems to test the positive correlation between the relative core size and robustness. Results are benchmarked with the comparison between connectance and robustness. The relative core size is shown to be a more reliable indicator of robustness, as this substructure property can better characterise the extent of redundancy in a food web.

Chapter 6 applies four most widely used centrality metrics to the 53 food webs and topologically important species are identified from different aspects. Targeted removal based on those centrality rankings are examined and related robustness are calculated.

Chapter 7 closes the thesis with a general conclusion on the findings and implications of the research, which is followed by an open discussion on the future directions.

Chapter 2

Data and Methods

Networks are often studied by referring properties at different scales. Local properties focus on the description of characteristics of individual nodes. Degree centrality quantifies the number of connection of a node is a metric that has been widely used [20, 73, 85, 91–95]. Betweenness centrality describes how many shortest paths pass through a node [61, 88, 96]. Global properties examine the behaviour of the whole network, such as the efficiency for flow [97], the integrity of the network in terms of stability and robustness [5, 17, 74, 75, 98]. In ecology, connectance which means the density of links in a network is often as seen an indicator of network robustness [38, 42, 77]. In addition, substructural properties can help reveal important cliques or subgroups in networks. For example, the core represents a densely connected subgraph that is said to be responsible for efficient information or energy exchange and linking all the periphery nodes [53, 54, 58, 99–101].

In this thesis, the primarily focus was to explore food web substructures by profiling a central and densely connected core-like structure [102] and examine how the substructures evolve in the face of external perturbations. In Chapter 3, the relative core size and core link density were used to examine the impact of drought on stream food web structures. In Chapter 4, the same core properties were studied between conventional and GMHT managed agricultural food webs to evaluate the impact of new management practice. In Chapter 5, the relative core size was examined to show if it is a strong indicator of food web robustness under node removal. The connectance was examined to quantify the impact of perturbations at the global scale in Chapter 3 to 5. Specifically, in Chapter 5, it was compared to the relative core size on their ability to characterise robustness. In Chapter 6, local properties including node centrality metrics were examined to assess the impact of different node removal sequence on food web robustness.

2.1 Data

The food webs analysed in this thesis consist of nodes as species and links as predation relationships. If species a is consumed by species b, the direction of the link is pointed from a to b. A resource represents the species that has no incoming links, and a consumer represents the species that feeds upon one or more species. In this thesis, food webs were treated as undirected networks when their substructural properties were studied. Food webs were only seen as directed networks when generating a spanning tree whereby the direction represents the way in which energy is transfered.

Food webs are often represented in a trophic structure shown in Figure 2.1, in which species are organised hierarchically to show how energy is transferred from the resource species to the top consumers. The resource species are located at the bottom level, and the consumers are located at a higher trophic levels.

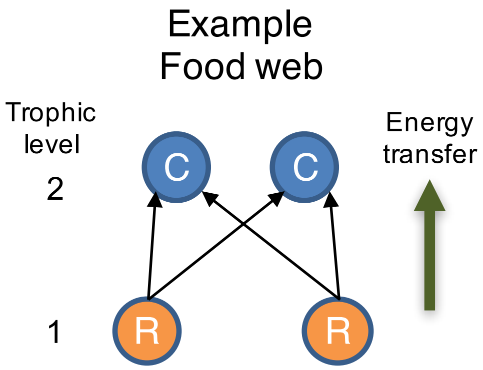


Figure 2.1: An example of a simple food web containing two resource species (labelled as R) and two consumer species (labelled as C) that feed upon the two resource species. Both the energy transfer direction and trophic level are labelled.

There are two main type of ecosystems from which food webs were sampled, namely the aquatic ecosystem and terrestrial ecosystem. Specifically, the eight food webs used in Chapter 3 were sampled from stream mesocosms under both control and drought condition (Data provided by Dr. Mark E. Ledger [36, 103]), and the links were obtained through the gut content analysis. The 502 food webs used in Chapter 4 were generated based on the data sampled from farmland across UK, which covers four types of crops (beet, maize, spring oilseed rape, and winter oilseed rape). The data was provided by Prof. David A. Bohan and Dr. Alireza Tamaddoni-Nezhad. The species interaction information was generated based on the Abductive Inductive Logic Programming (A/ILP), and the detail of how the method is developed is described in [104–106]. The 53 food webs used in Chapter 5 and 6 are all from water-based ecosystems, which covers marine, freshwater, and lentic. The classic food webs including Benguela, Broadstone stream, Skipwith Pond, Ythan Estuary, and Tuesday Lakes were from the Cheddar Package [107]. The 20 highly resolved stream food webs over a wide pH range was provided by Dr. K. Layer [81, 108]. The 18 US riverine food webs were provided by Dr. C. Mulder. The Kennet food web was provided by Prof. Guy Woodward and Dr. Clare Gray [109]. For the details of each food web, information including the ecosystem site name, species identity, and related literatures, please refer to the Data set section in each Chapter.

2.2 Food web analysis metrics

A food web with S species and L links is represented as $G = (S, L)$. A_{ij} represents the adjacency matrix of graph G. G is treated as an un-weighted network if $a_{ij} = 1$ means there is a link between node n_i and n_j and $a_{ij} = 0$ otherwise. Topological based properties are extensively used in food web studies, from the simplest metrics such as the total number of nodes (S) or links (L) in a food web, towards some summarised metrics such as the fraction of basal (B) / intermediate (I) / top-level (T) species, and the number of links per species (L/S) [38, 44, 81, 110]. L/S also reflects food web complexity, as it calculates the average number of interactions a species may have [111]. Connectance, as the most widely used metric, measures the probability that a link exists between any two selected species [38, 42, 77] (Eq. 2.1).

$$C = \frac{L}{S^2} \tag{2.1}$$

Connectance normally ranges from 0.06 to 0.31 [6, 15, 42], whereby theoretically it spans between 0 to 1 [43]. In reality, food webs always contain resource species which are located at the bottom level of a trophic structure and they do not fed upon each other. Therefore, the range of connectance is narrowed. Specifically, marine food webs [80] normally have lower connectance than freshwater food webs, which results from marine webs in general have more trophic levels than freshwater food webs and are likely to encounter fewer and less severe perturbations than freshwater ecosystems [36, 50].

2.3 Substructures

Recently, core/periphery profiling is gaining increasing attention as this structure can reflect network properties that are not able to be detected using macroscopic or microscopic network metrics. A working definition of core/periphery structure is that the core contains all the central nodes and links while periphery nodes have links mostly with the core (Figure 2.2) [54].

There exist various methods to identify whether a network contains a core/periphery structure [52, 54, 57, 64, 99, 101, 112]. The method presented in [112] classified networks into different subclasses based on the network spectrum, which can be applied to various types of empirical networks. Specifically, most ecological networks have been classified as a homogeneous network which contains no *bottlenecks*, and therefore, ecological networks tend to exhibit higher robustness when compared to other types of empirical networks [113]. *Bottlenecks* are high centrality species which are responsible to connect the rest of the network together, therefore, the removal of which would cause devastating effect to the network [113]. The heterogeneity of complex networks reflects the uniformity or dissimilarity in the structure, i.e., the minority of nodes have high degree whilst the majority of the nodes have very low degree [15, 114, 115].

Various types of core/periphery structures demonstrate the complex interaction patterns among nodes, indicating the heterogeneity properties of the network. A typical portrait of core/periphery structure is shown in Figure 2.2. Real-world examples include: the core of the London underground network contains King's Cross and Waterloo Station which are both located in central London and being major train station as well [57]. The core of the brain network (nodes represent cortical or subcortical areas and links represent whether they are functional related or not) contains regions that are most active in information change during learning process whilst being the most stable regions when learning tasks are changed [58]. Examples on what are the core structures in real world network and their implications are summarised in Appendix B.2. In summary, the key function of the core is to govern most of the energy or information flow within the network, as well as being stable both across time scale and different environment scenarios.

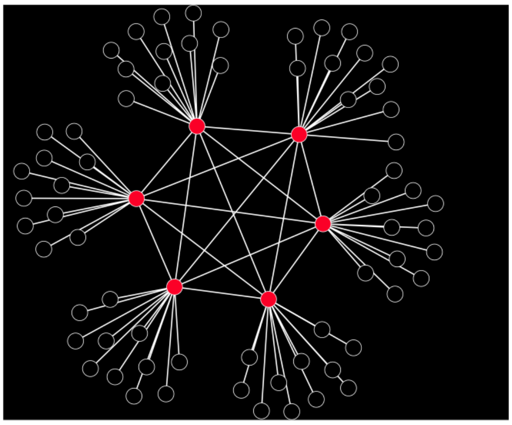


Figure 2.2: A typical example of a core (formed by red nodes)/periphery (formed by the rest of black nodes) structure. Red nodes also formed the rich-club in the network.

Since the first proposition of the existence and topological importance of a core/periphery structure in social networks [99], it has attracted continuously attention with various core detection algorithms being developed for different focuses [57, 58, 64, 101]. For the classical core/periphery partition method [99], it lacks a consistent quantifiable method to prove what one suspects is the core is indeed the core. Later proposed method uses *core-periphery coefficient* [64] to compare the closeness within the real network core and that in random networks, which provides statistical significance to how well the core is defined. However, this core-periphery coefficient only emphasises whether the core has the similar link density as a clique while neglects the connectivity between the core and the rest of the nodes [64]. A recently developed method using *Core score* [57], which ranges from 0 to 1, to decide how well each node is qualified to be included in the core. However, there are arbitrary parameters which are used to define the number of core/periphery nodes and how well the partition to be expected, increases the uncertainty and the complexity of this method. It requires long computational time and it is not feasible for large-scale networks [57]. A more recent core profiling method significantly reduces the computational complexity by using pre-defined node ranking according to a certain node weight (e.g., degree) [102]. The core is defined by computing each node's connectivity with nodes of a higher rank, enabling the comparison between the core and the rest of the network, therefore eliminates the introducing of random networks [102]. Moreover, the core, which is profiled by this method, emphasises not only its central position within the network (all the nodes in the core have significantly shorter paths to every other node), but also its densely interconnected feature (the core is necessarily a rich-club, defined in [116]).

2.3.1 Rich-core profiling

The *rich-core* profiling method exploited in this thesis incorporates the rich-club coefficient and node connectivity behaviour to detect a core/periphery structure within a network [102]. In a undirected and unweighted graph G, nodes are ranked based on the degree. For each node r, k_r represents its degree, k_r^+ represents how many nodes link with r that have a higher degree than r. For a network with N nodes, a resulting k_r^+ sequence is generated with length N. The core contains nodes whose $k_r^{*+} > {k_r}^+$ if $r > r^*$. That is, the boundary of the core is detected at the point after where k_r^+ reaches its peak and decreases afterwards. The advantage of applying this method to detect the core is twofold. First, a pre-ranked node sequence makes the computation faster for large networks. Second, the envelop of the k_r^+ sequence is able to explicitly show the transition point where the boundary of the core is defined. The core in the context of the thesis refers to a densely connected clique among high degree nodes [102], and this two-class partition method is always able to identify a core and a periphery in non-random networks.

This method is also applicable on directed or (and) weighted networks. For a directed network, node ranking r is defined by either in-degree or out-degree of each node. k_r^+ no longer represents how many links node r has with a higher degree, but represents how many incoming/outgoing links r has with a higher degree node. Therefore, the implication of the core becomes the original source of the information spreading or the terminal of the information receiving substructure. For a weighted network, node ranking r is defined as the sum of weight from all its connected link, which is used to calculate the k_r^+ value. Therefore, the core not only represents the most connectivity in the network, but also indicates most of the energy or information flow within and towards it.

A simple example of a synthetic food web under control condition is shown in Figure 2.3**a**, containing basal level ('B') species, intermediate level ('I') species, and the top-level ('T') species. In order the calculate the k_r^+ metric, species are first ranked in the descending order of degree, shown on the x-axis of Figure 2.3**b**. The y-axis of Figure 2.3**b** indicates the number of interactions that particular node with those nodes of a higher degree than itself. This gives an idea of the local connectivity among high ranking nodes. A synthetic food web under drought condition (where large top predators 'T1', 'T2', 'T3' and 'T4' have gone extinct compared to the web in Figure 2.3**a**) is shown in Figure 2.3**c**, and its core profiling is shown in Figure 2.3**d**. The core (peak value k_r^+) is at node 'T3' in Figure 2.3**b**, and node 'B4' in Figure 2.3**d**.

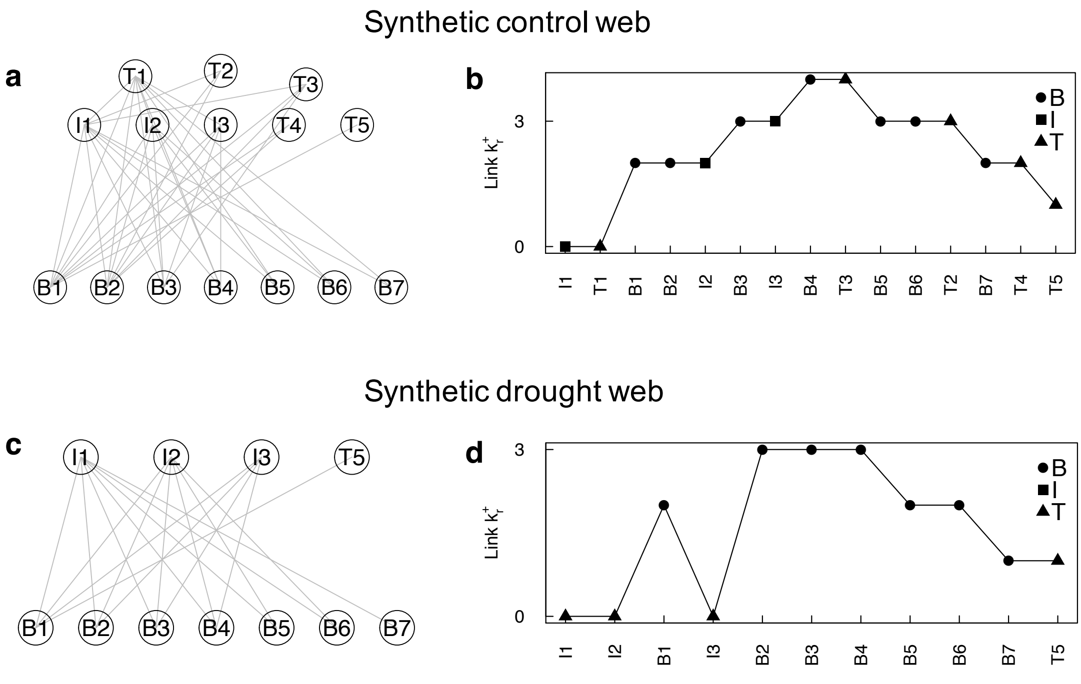


Figure 2.3: **a** - Synthetic food web under control condition. **b** - Core profiling of the control web **a**. **c** - Synthetic food web under drought condition. **d** - Core profiling of the drought web **c**. **B** (circle) represents basal species, **I** (square) represents intermediate-level species, and **T** (triangle) represents top-level species.

2.3.2 The rich-club coefficient

The Rich-club coefficient measures the density of connections (explained below) between the high degree nodes [116]. It was first proposed to describe the connectivity pattern among high degree nodes of Internet at the Autonomous System (AS) level: an extremely large proportion of links are distributed within a small fraction of very high degree nodes, forming a richly connected core. This rich core is able to maintain its performance and connectivity under various perturbations due to its densely connected behaviour, therefore helps maintain the whole network robustness. This rich core substructure also ensures a high efficiency information flow from the central nodes to the rest of the network.

Rich clubs have been identified in many other types of networks as well, such as power grid networks [117], transportation networks [118], scientific collaboration networks [117, 118], and brain networks [119]. The rich-club structure in the aforementioned networks contains important nodes that are both with high degree properties and also connecting to all the other low degree nodes in order to transfer energy, accessibility, information, or exert control.

The quantitative representation of the rich-club coefficient is (ϕ_r) is defined as [120]:

$$\phi_r = \frac{2}{r(r-1)} \sum_{i=1}^{r} \Delta E_i = \frac{2E_r}{r(r-1)}, \tag{2.2}$$

where r is rank of the descending order of the node degree, E_r is the number of links shared by the higher ranked nodes, and $r(r-1)/2$ is the maximum number of links that these nodes can share. The connectivity of a subnetwork consisting of chosen nodes is given by ϕ_{r^*}, whereby a fully connected subnetwork has a value of $\phi_{r^*} = 1$ and a fully disconnected subnetwork gives $\phi_{r^*} = 0$.

2.3.3 Food web redundancy

Redundancy, in general, is defined as two or more components that share the same functional role within a system, such that the malfunctioning of one of these components will not affect the services provided by the whole system [121]. In ecology, it was claimed that the robustness of ecosystem also relies upon certain level of redundancy to prevent cascade species extinctions under disturbances [4, 81, 122–124]. The level of redundancy was quantified by the number of species that are independent to each other (i.e., there is no predation relationship) and share the same functional role within an ecosystem, for example, there are two or more resource species consumed by the same consumer, therefore, the extinction of any one of the resource species will not cause further secondary extinction towards consumers [12, 82, 121, 125, 126]. Later studies also shown that food webs in general are densely connected and the intertwined pathways can provide redundancy [54, 99], which is able to prevent potential secondary extinction when various perturbation happens, such as environmental warming [122], acidification [81, 123] and drought [4, 124]. In Chapter 3 and 4, redundancy was interpreted as the alternative paths within the core, whereby the relative size of the core indicates the food web robustness. In Chapter 5, substructure redundancy was introduced based on the core/periphery partition of food webs. The core redundancy was defined as the proportion of the alternative paths within the core against the total number of alternative paths in the whole food web.

2.4 Food web robustness measurement

In network science, a common way to assess network robustness is to gauge the rate in which a network collapses after a proportion of nodes being removed, and metrics used to gauge the integrity of a network include efficiency [97] and the size of the giant component [127]. For example, the robustness (R_x) among a set of networks is compared by measuring the size of the giant component when x% of the nodes are removed [128], where a giant component is the largest connected component of a network [92, 129–131].

Recently there has been a focus on quantifying food web robustness based on simulated species removal aiming to estimate the vulnerability of existing ecosystem to possible perturbations [17, 60, 80, 82, 132–135]. The simulated species removal is either based on randomly selection or certain node ranking methods [17, 80, 82]. Food webs are remarkably robust to realistic extinction scenarios (i.e., those species with limited geographical distribution are likely to go extinct [133]), as opposed to the non-ecologically informed scenarios usually employed (i.e., the random species loss), suggesting that food web has the ability to adapt thereby minimises the negative consequences of realistic species extinction. Ultimately the goal is often to investigate food web robustness under different environmental perturbations as a way to evaluate the consequences to both the structure and functions, which helps target conservation efforts toward those species and ecosystems that most need it [50, 81, 134–136].

In ecology, robustness is commonly referred to R_{50} which quantifies the amount of secondary extinction [17] The measurement is described in the following way: 1) At each iteration, one species is selected and is removed from the network, together with its associated links. 2) The effect of species removal propagates towards the upper levels of the network, and 'bottom-up' effects [5, 17, 74, 75] (i.e., predator lost all its prey), if there is any, will be evaluated. 3) If this procedure causes any higher level species to lose all its prey, secondary extinction will be triggered. The effect of any secondary extinction is consecutively evaluated until no further extinction would happen within this iteration. 4) Procedure 1 will repeat until the entire food web becomes half of the original network size.

Simulated species removal can be based on different node ranking sequences. First, perturbation is simulated by randomly removing species across the network [17]. Second, perturbation is simulated by removing species with the highest degree, which represents the worst case scenario [82]. Removal under the re-ranked procedure was claimed to be able to achieve the most efficiency [71]. Therefore, before each targeted removal, nodes were re-ranked based on the centrality method selected. Throughout the simulation, neither isolated nodes nor predators that lose all its prey would exist in the web.

Finally, robustness is quantified as the amount of primary extinctions (Q) in order to generate 50% of the total species loss:

$$R_{50} = \frac{Q}{N}, \tag{2.3}$$

where N is the number of nodes in the original web.

This definition gives a maximum possible robustness of 0.5 (all extinctions come from Q removals) and a minimum of $1/N$ ($1/N$ happens if 1 removal causes $\leq 0.5N - 1$ secondary extinctions). The cumulative secondary extinction caused by sequential removal is usually represented as Figure 3.8 or Figure 6.6. The greater the skew, the more secondary extinctions occur and the less robust the food web will be.

The difference of R_{50} and the robustness measurement in general (R_x) lies in the stopping point of the node removing iteration, i.e., when the certain remaining size of the network is reached, or when the certain proportion of nodes is removed. Yet, based on a recent study on the reliability of R_x applied to food webs where extinction threshold x ranged from 5% to 95%, the relative proportion of secondary extinction did not show significant changes when different proportion of species was manually removed [98]. Also, when the collapse threshold was restricted within $50\% \pm 20\%$, R_x was found to be relatively invariant. On the other hand, R_{50} has been shown to be a reliable measure of food web robustness [17] which has been used in this thesis.

However, if the focus is to evaluate the importance of removing certain species in a single food web, R_{50} exhibits two major limitations. First, it measures only binary networks which neglected the effects of species with different abundance. Second, the procedure is only based on the original snapshots of food webs which ignores the fact that ecosystems are able to evolve according to perturbations. Later studies expanded this binary model from different perspectives. Robustness model based on weighted networks incorporates species abundance as a factor, which was proven to be able to capture more realistic responses of natural ecosystem, i.e., fewer secondary extinctions are observed when removing specialised species with a relatively low abundance [82, 135, 137]. Multi-species removal is used to simulate habitat removal as a consequence of regime shifts and human activities, i.e., the loss of a group of species that live in a certain habitat. This can be applied to model the robustness of ecosystems which contain top predators that depend heavily on species mostly from a single habitat [84, 138].

2.5 Statistical evaluation methodology

2.5.1 Correlation (Pearson and Spearman)

Correlation is used to quantify the strength of association between two variables. For the continuous and normally distributed numerical data, the Pearson correlation coefficient (r) was used, while for the discrete ranked data or non-normal distributed data, the Spearman correlation coefficient (ρ) was used. While r measures the strength of the two data having a positive or negative linear relationship, ρ measures whether the two data would increase or decrease monotonically with each other.

Suppose there are one dataset $X = x_1, ..., x_n$ and another dataset with the same length of X, $Y = y_1, ..., y_n$, the Pearson's correlation coefficient r is defined as:

$$r = \frac{\mathrm{cov}(X, Y)}{\sigma_X \sigma_Y} \tag{2.4}$$

where cov is the covariance, and σ_X is the standard deviation of dataset X.

The Spearman correlation coefficient ρ is defined as:

$$\rho = \frac{\mathrm{cov}(rg_X, rg_Y)}{\sigma_{rg_X} \sigma_{rg_Y}} \tag{2.5}$$

where $\mathrm{cov}(rg_X, rg_Y)$ is the covariance of the ranked data, and σ_{rg_X} is the standard deviation of the ranked data rg_X.

In Chapter 4, the Pearson correlation was used to compare food web size between conventional treatment and that from the new treatment. It was also applied to measure the strength of correlation between food web robustness and the relative core size or connectance in Chapter 5. In Chapter 6, the Pearson correlation method was used in the same way as in Chapter 5, whilst food web robustness was also compared to other properties such as the proportion of basal/intermediate/top level species in additional to the relative core size and connectance. The Spearman correlation was applied in Chapter 6 to quantify the similarity among centrality ranking indices.

2.5.2 Null model

In network science, random networks are modelled with pre-defined topological features that maintain a certain degree of randomisation among replicates [139]. This type of null model conserves the total number of nodes and links compared to the original network, as well as the degree of each node. The rewiring procedure ensures that all the neighbors of each node is randomly assigned, therefore, the connectivity pattern of the network is changed. By comparing an empirical web with its randomised networks, it is possible to determine whether the pattern (i.e., assortativity, rich-club phenomenon) observed in the real network would happen by chance (if no difference from the random network) or are statistically significant (if huge difference from the random network).

This null model has applications in many areas. For example, Internets or World-wide-webs are examples of "scale-free" networks that show a power-law degree distribution [140, 141] whilst food webs, on the other hand, have uniform or exponential degree distributions [40, 42]. Those examples make a random graph with a Poisson degree distribution not representable for real-world networks. In order to preserve the important topological properties of the real networks while producing the appropriate random networks, a null model is needed to generate corresponding random networks.

Through this thesis, null model was used to identify non-random topological features of real networks by comparing them with random re-wired networks [139]. When applied to food webs, random webs generated by this null model preserve properties of empirical webs while also have links that are not able to be observed in reality. The more improbable a configuration of links is, the more *re-organisation* is required to be in place to do so [44, 142, 143]. Therefore, important food web structural patterns can be unveiled.

2.5.3 Z-score

Z-score (standard score) is the signed standard deviation (σ) between the actual data and the mean [144]. If the actual data is larger than the mean, the z-score has a positive sign, otherwise, it has a negative sign. In general, the z-score is defined as:

$$z = \frac{x - \mu}{\sigma}, \tag{2.6}$$

where μ is the mean of the null model [145].

In order to use z-score to testify the null hypothesis on the level of significance, the critical value of the confidence level is used. For example, a 95% confidence level (i.e., p-value > 0.05) is associated with a critical z-score within ±1.96. Z-score higher than 1.96 or lower than -1.96 indicates the empirical data has the significant difference with the null model [144].

2.5.4 Bray-Curtis dissimilarity

In ecology, Bray-Curtis dissimilarity is used to assess how taxon composition varies between two different ecosystems based on the count of each species. The index is given by:

$$b_{ii'} = \frac{\sum_{j=1}^{J} |n_{ij} - n_{i'j}|}{\sum_{j=1}^{J} n_{ij}}, \tag{2.7}$$

where i and i' correspond to two comparing sites, n_{ij} and $n_{i'j}$ refer to the species count for species j. (J refers to the union of all the species) in the two sites of interest. The parameter b varies between 0 and 1, where 1 indicates the highest variability between two sites, i.e., no common species found in both sites, and 0 indicates the lowest variability, i.e., species are identical in identity and count number between both sites.

2.6 Spanning tree

In mathematics, a *tree* is an undirected graph in which any two nodes can only be linked with a single link. In other words, a tree is a type of graph with no loops and the removal of any single link would break down the graph into disconnected components [146]. A spanning tree is a subgraph in the form of a tree of the original graph, which connects all the nodes with the minimum number of links.

In this thesis, the construction of a spanning tree structure is based on the above definition. Here, the construction of a spanning tree takes into consideration the direction of the predation relationships. In order to fulfil the requirement that 1) the spanning tree contains minimal number of links such that the removal of any link would result in isolated species or further secondary extinctions, and 2) each predator can be fed upon one and only one prey; an additional node representing the environment that provides energy to all the resource species was added to make sure the spanning tree of a food web is a connected graph. Specifically, if Figure 2.4**a** represents a simple example of a food web, Figure 2.4**b** represents all the possible constructions of the spanning tree structure based on the original food web [12]. This type of reconstruction was applied to empirical food webs in Chapter 5.

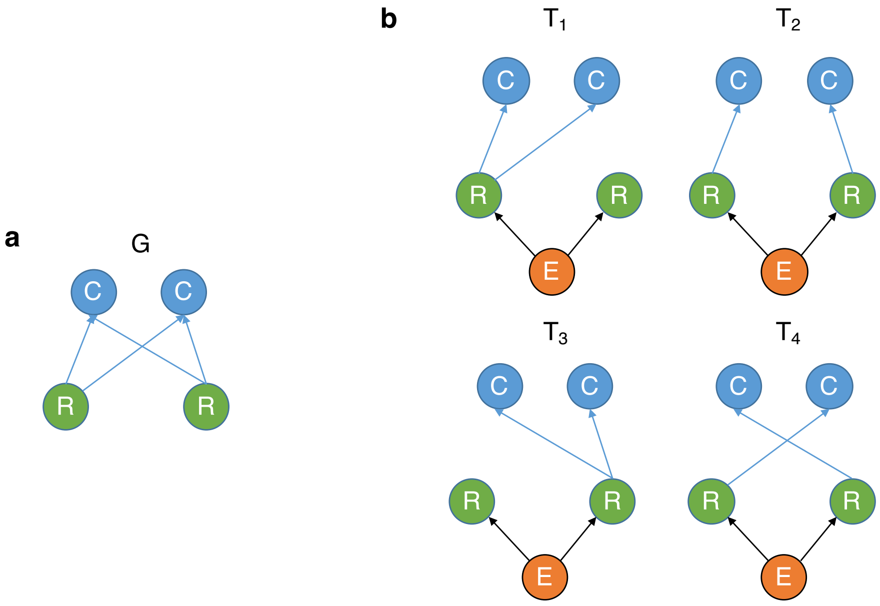


Figure 2.4: **a** - Original food web G with two resource species (R, green) and two consumer species (R, blue). **b** - Four possible construction of the spanning tree structures (T_1 to T_4) based on G in **a**. Node E (orange) is the manually added node representing the environment that provides energy to all the resource species. All the nodes are linked with only one link. Each spanning tree structure represents a possible type of energy flow from resource to consumers.

2.7 Centrality indices

Centrality aims to characterise which nodes are important and thus have a "central" position in the network based on the network structure. Here only centrality metrics that are linked to the robustness properties of ecological networks are selected and studied. Four well-known and widely applied centrality indices are summarised, including the node-level degree centrality, the substructure-level betweenness centrality, and the network-level closeness and eigenvector centrality.

- Degree centrality

The degree centrality is the simplest and most frequently used metric to quantify the importance of nodes that maintain the integrity of the network structure [20, 73, 85, 91–95]. The degree of a node is the number of links that connects to it and it is defined as

$$k_i = \sum_{j \in N} a_{ij} \tag{2.8}$$

where a_{ij} is given in the adjacency matrix, and N is the number of nodes in the network.

The degree is a purely local centrality measure, as it considers only the number (or weight) of the connections of the node to its connected neighbours. To obtain a more macroscopic perspective, it is often insightful to look at the degree distribution, which summarises the frequencies with which different degrees appear in the network. For example, a power-law degree distribution enables a network to be robust against random failures but vulnerable to attacks on high degree nodes [94, 95].

In ecology, high degree species are found with higher probability to be present in more fragments (e.g., geometry niches) within a certain area, which is named *nestedness* from the view of geographic occurrence pattern [147]. Species degree also reflects the survivability of this species, therefore helps with the conservation purposes, the degree change of a species could be consequences of climate change and could help predict its survivability [20, 85, 148, 149].

- Betweenness centrality

Betweenness centrality counts the fraction of shortest paths going through a given node [61, 88, 96], and it is defined as:

$$C_B = \frac{\sum_{s \neq n \neq t} \sigma_{st}(n)/\sigma_{st}}{(N-1)(N-2)/2} \tag{2.9}$$

where σ_{st} is the total number of shortest paths from node s to node t and $\sigma_{st}(n)$ is the number of shortest paths σ_{st} passing through node n. Thus betweenness can be seen as a substructure level metrics.

In biology, nodes (or links) with the highest betweenness are often claimed to be key nodes (or links) that have multi-functional properties [64, 89, 150–153]. In ecology, the links with highest edge betweenness are also often considered as a bridge between different communities within a single food web [61, 154]. For example, Girvan & Newman proved the removal of the highest betweenness links in Chesapeak Bay food web resulted in the splitting of pelagic and benthic sub-communities [61]. Food webs that lack high betweenness nodes, which is named as "good expansion networks", are more robust to perturbations [19].

- Closeness centrality

Closeness centrality calculates the "distance" (d_{ij}) given by the average shortest path between a given node i and all other nodes in the network [150], and it is defined as [155]:

$$C_C = \frac{N-1}{\sum_{j \in G} d_{ij}} \tag{2.10}$$

It makes closeness centrality a macroscale network metrics. This measure is meaningful only for connected graphs, otherwise $d_{ij} = +\infty$ when there is no path between i and j.

In food webs, species with a high closeness centrality are considered to spread (or accumulating) energy faster within the network and are therefore deemed to be of higher importance [20, 64, 85, 150, 154]. Specifically, in host-parasite networks, the parasites with the highest closeness centrality consume the widest range of prey. Those parasites form a highly connected central substructure in the network which is robust to external perturbations [154, 156].

- Eigenvector centrality

Eigenvector centrality measures the level of influence of a node to all the other nodes in the network. The eigenvector centrality v_i of node i is defined as [155] :

$$v_i = \lambda^{-1} \sum\nolimits_j a_{ij} v_j \tag{2.11}$$

where v_i is the score for node i, λ is a constant and the sum is over all the nodes. Eq. 2.11 is expressed as the eigen–problem:

$$\hat{\mathbf{a}}\mathbf{v} = \lambda \mathbf{v} \tag{2.12}$$

The eigenvector centrality v_i is given by the i entry of the eigenvector $\mathbf{v}$ corresponding to the largest eigenvalue λ.

A node with a high eigenvector centrality shares most of its connections with other important nodes [87, 157, 158]. It can be seen as a globalised version of degree centrality, as the importance is not localised around the neighbors of this nodes, but considers how its neighbors would locate in the network. For example, a node connecting to all the high degree nodes has higher eigenvector centrality than the same-degree-node that connects to low degree nodes. High eigenvector centrality nodes are further classified into two categories: *community core* nodes (the removal of which would make the network fall into many small isolated sub-communities) and *bridge* nodes (the removal of which would separate different communities more clearly) [159] (Figure 2.5). If an attack happens, the community's core nodes would be more important as the removal of any would lead to many isolated species leading to further secondary extinctions [4, 5, 20, 85, 148, 149].

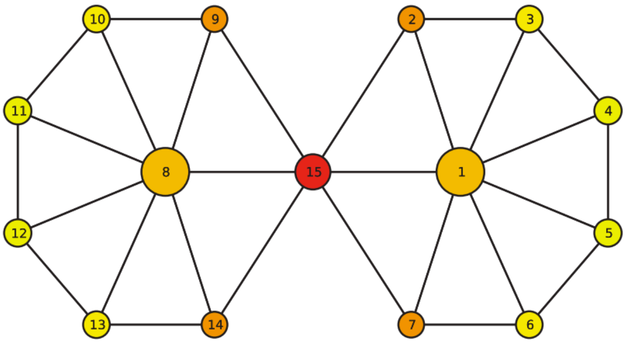


Figure 2.5: Node 1 & 8 are the community core nodes. Node 15 is the bridge core node. The removal of node 1 & 8 will destroy the community structure and decrease the efficiency of each node to reach every other node. The removal of node 15 will separate two sub-communities while the connectivity within each sub-community remains the same [159].

- Reliability of the node-ranking metrics

The effectiveness of targeted attack depends on the accurate measure of the aforementioned centrality measures. In reality, a complete view of the topology of a certain network is usually not available, for example, there exists missing species and its associated predation links in almost all the food webs due to the sampling error. Thus it is important to decide whether the centrality measures are able to return the desired results with imperfect data [93]. Research in [93] has shown that the ranking accuracy of the four centrality metrics listed above are all robust to errors up to 25% of missing data, and this was shown using in random networks with 100 nodes, and the probability that a link exists between any two nodes equals to 0.5 (i.e., connectance roughly equals to 0.245).

From another perspective, the aforementioned network centrality metrics have been proved to generate node ranking sequences with high similarity within different types of networks, for example, in scale-free networks, the closeness ranking shares the most similarity with eigenvector ranking, while degree shares the most similarity with betweenness

[128, 160]. A recent study highlighted the effectiveness of betweenness-based attack on a set of random scale-free networks as well as a few real world networks that follow a scale-free degree distribution [153]. However, its effectiveness on other types of networks, for example, food webs, are still unknown. In Chapter 6, those centrality metrics were compared by analysing a set of empirical food webs, and their similarity in node ranking will also be compared.

Chapter 3

Effect of drought on food web substructures

3.1 Research background and overview

Drought is a major climate issue, and it is predicted to become more severe within the next few decades [161]. Drought caused by intermittency in streams [162], declining precipitation [163], and the extensive human activities [161, 164] result in changes in both species composition and biodiversity in freshwater ecosystems. The consequences on food webs from drought can be seen from two perspectives: whether species can survive under drought; and if so, how will it adapt to the new environment [165]. Previous studies showed that the response of individual species under drought (i.e., species' presence or absence, the change of species biomass production, etc.) varies significantly. In addition, there exists synecological research on how network level properties, such as connectance would fluctuate under climate change [166, 167], however, changes of species interactions under different environmental stressors remain unclear. In network science, the presence of a core/periphery structure is found to be responsible for maintaining the network structure and faciliating flows in various kinds of non-ecological networks [53, 54, 58, 99–

101]. The significance of presenting a core/periphery structure for food web responses to an environmental perturbation - drought - is reported here for the first time.

In this chapter, network analysis methods were applied to provide a novel way to examine the consequences of environmental perturbations on food webs. The previously unexplored substructural changes were assessed within the food webs by comparing between food webs under control and drought conditions. For each food web, the rich-core profiling technique was applied to detect if food webs have a clearly divided core/periphery substructure [102]. Next, the link density within the core [116] was also measured. The random networks were generated through the degree-preserving null model in which links were random rewired [139]. The null model was used to show if there exist unique patterns of species connectivity in empirical food webs compared to random networks. Species extinction was either originated from core or periphery, and the results can help identify the difference in substructure fragility. Species movement between core and periphery was also examined as it is an indicator of network adaptability. Finally, whether the presence of a core structure would contribute to the maintenance of food web robustness was examined via simulating species loss *in silico* both randomly and strategically in descending order of degree.

3.2 Methods

3.2.1 Data set

The set of food webs used in this chapter comprises of four pairs of food webs sampled from stream mesocosm under control and drought conditions (Data provided by Dr. Mark E. Ledger). The drought treatment was applied to the streams for two years, and species were sampled at the end of the treatment. The interactions between species were generated through gut content analysis (involving 3643 individual species) [50]. The eight food webs are among the most highly resolved to date, comprising 783 pairwise trophic interactions and 74 trophic elements in total.

The eight mesocosm food webs are realistic representation of natural systems regarding to their network complexity (i.e., number of species and their overall link density) and their structural similarity by comparing to 82 natural river food webs [168]. Within each of the individual food web, species were categorised into three trophic levels: basal (B), intermediate (I) and top (T). A basal species was defined as a species with no prey; a top-level species was referred to as a species with no predators; and the rest were defined as intermediate species.

3.2.2 Core/periphery profiling and the rich-club coefficient

The details of how to profile a core is described in Methods, section 2.3.1. To summarise its application to food webs: First, the nodes (representing the species) are ranked in the decreasing order of their degree. For each node, the number of links this node shares with nodes of higher degree (represented as k_r^+) was recorded, where r is the rank of a node. The highest degree node always has $k_r^+ = 0$. In theory, the maximum k_r^+ for each node r is $r - 1$, if this node links with all the nodes with a higher rank. Along this k_r^+ profile, the peak r^* was detected if k_r^{*+} is always larger than k_r^+ if $r > r^*$. All the nodes before the peak were grouped in the core substructure and the rest to be in the periphery.

Species turnover between substructures was analysed in two ways. The core (or periphery) species in each of the control webs were compared with the species in its corresponding drought web to get: 1) in which substructure (i.e., core or periphery) of the control food web did the species extinction mostly happen; and 2) the proportion of species movement from the core of control webs to the periphery of drought webs, and vice versa.

The density of interactions within the core and across the web was measured using the "rich-club" coefficient, ϕ_r, where r represents the ranking of the nodes, and ϕ_r measures the actual number of links existed among selected nodes divided by the maximum possible number of links [116] (section 2.3.2). The sequence of ϕ_r from the highest degree node ($r = 1$) to the lowest degree node indicates changes of the subgroup (contains all the nodes with rank $> r$) connectivity. A fully connected subgraph gives $\phi_r = 1$. When more lower degree nodes being included in the calculation, ϕ_r will decrease after reaching its maximum value. The last value in the sequence of ϕ_r measures the density of the whole network.

3.2.3 Null model comparison

The construction of an ensemble of null models is achieved by randomly rewiring links among nodes [139]. The link rewiring procedure conserves the degree of each node, therefore the overall degree distribution is unchanged [139]. For each of the random networks, $\phi_{r_{null}}$ is calculated and is compared with that of the empirical network, ϕ_r. Z-score (Section 2.5.3) is a statistical metric that measures how far the actual value deviates from the corresponding average ϕ_r among the null models (i.e., a z-score greater/lower than 0 indicates the actual value is above/below the mean). The analysis was able to benchmark how far the real food web assemblage pattern deviates from the random networks, which gauges the level of organisation of predation links triggered by drought.

3.2.4 Statistical tests

One-tailed t-test was used to examine whether the proportion of extinction species in the core and periphery, and whether the proportion of species movement from core to periphery and vice versa showed any significant differences. Two tailed t-test was used to examine whether the relative core size of control and drought webs, and whether the robustness of control and drought webs showed any significant differences (Table 3-A).

Table 3-A: Summary of two independent samples t-tests.

What has been tested?	**Dependent variables**[*]	**Independent variables**	H_0 [†]
Has the relative core size changed in response to drought?	Relative core size	Treatment with two levels (control and drought)	$\mu \neq \mu_0$
Is species extinction greater in the periphery than in core?	% of species extinction	Substructure with two levels (core and periphery)	$\mu > \mu_0$
Do more species move from core to periphery than vice versa?	% of species movement	Substructure with two levels (core and periphery)	$\mu < \mu_0$
Are control webs more robust than drought ones under random removal?	Robustness	Treatment with two levels (control and drought)	$\mu \neq \mu_0$
Are control webs more robust than drought ones under targeted removal?	Robustness	Treatment with two levels (control and drought)	$\mu \neq \mu_0$

[*] The original data in all comparisons were examined using Shapiro-Wilk test and they all met the normality assumption. Data were on proportions and therefore arcsine transformation was applied. The normality assumption held after transformation.

[†] The null hypothesis H_0 being $\mu \neq \mu_0$ indicates a two-tailed t-test, while $\mu > \mu_0$ or $\mu < \mu_0$ indicates a one-tailed t-test. μ_0 represents the mean of variables related to the core or the control webs, while μ represents the mean of variables related to the periphery or the drought webs.

3.2.5 Food web robustness

Food web robustness was studied by simulating nodes removal *in silico*, and the proportion of secondary extinction against original web size (when species loss has reached 50% the total species) was measured [4, 16, 17, 42, 78] (Section 2.4). A secondary extinction happens if a predator loses all its prey (i.e., a node loses all its incoming links). This robustness analysis provides estimates of how vulnerable the ecosystem would be when facing species extinctions. The more iterations of primary species removal a web needs to achieve a certain amount of secondary extinction, the higher robustness a web would have. Different node removal strategies may have different implications when characterising robustness. For example, high degree node removal (i.e., the highest degree species is always the one to be removed in each iteration) is able to simulate a 'worst-case' scenario on the consequences of species extinction while random removal (each iteration, a species is randomly selected) simulates how species would behave against random perturbations [82, 169].

3.3 The impact of drought on food web substructures

3.3.1 Cores in food webs

A clear core/periphery structure was shown in all eight food webs (Figure 3.1), with a distinct peak in their core profiles (indicated by a vertical line, at which the number of links k_r^+ is at its maximum, and after which the k_r^+ is always smaller than the peak). The food web cores contained species from all trophic levels. The absolute core size is always smaller in drought webs than that in control webs, as the web size is always smaller under drought. However, food web cores are accounted for (on average) 50% of the species in both control and drought webs (Table 3-B). The relative core size was unchanged by drought (t-test on arcsine transformed proportion data, d.f.=3, p=0.16; Table 3-C), despite absolute species losses of 25%.

Table 3-B: Summary on properties related to the core. The core size, species extinction (E) found in the core and periphery, and the species re-alignment between the two regions when comparing a control web (C) with its respective drought web (D).

Web pair	Number of species		Core size (% of web size)		Number of species lost from core		Number of species lost from periphery	
	C	D	C	D	to E	to periphery	to E	to core
1	59	47	30 (50%)	27(57%)	4(13%)	4(13%)	13(45%)	4(14%)
2	63	46	31 (49%)	20(43%)	8(26%)	7(23%)	12(38%)	3(9%)
3	61	49	36 (59%)	23(46%)	8(22%)	7(19%)	9(36%)	0(0%)
4	65	52	38 (58%)	22(42%)	7(18%)	11(29%)	8(30%)	0(0%)

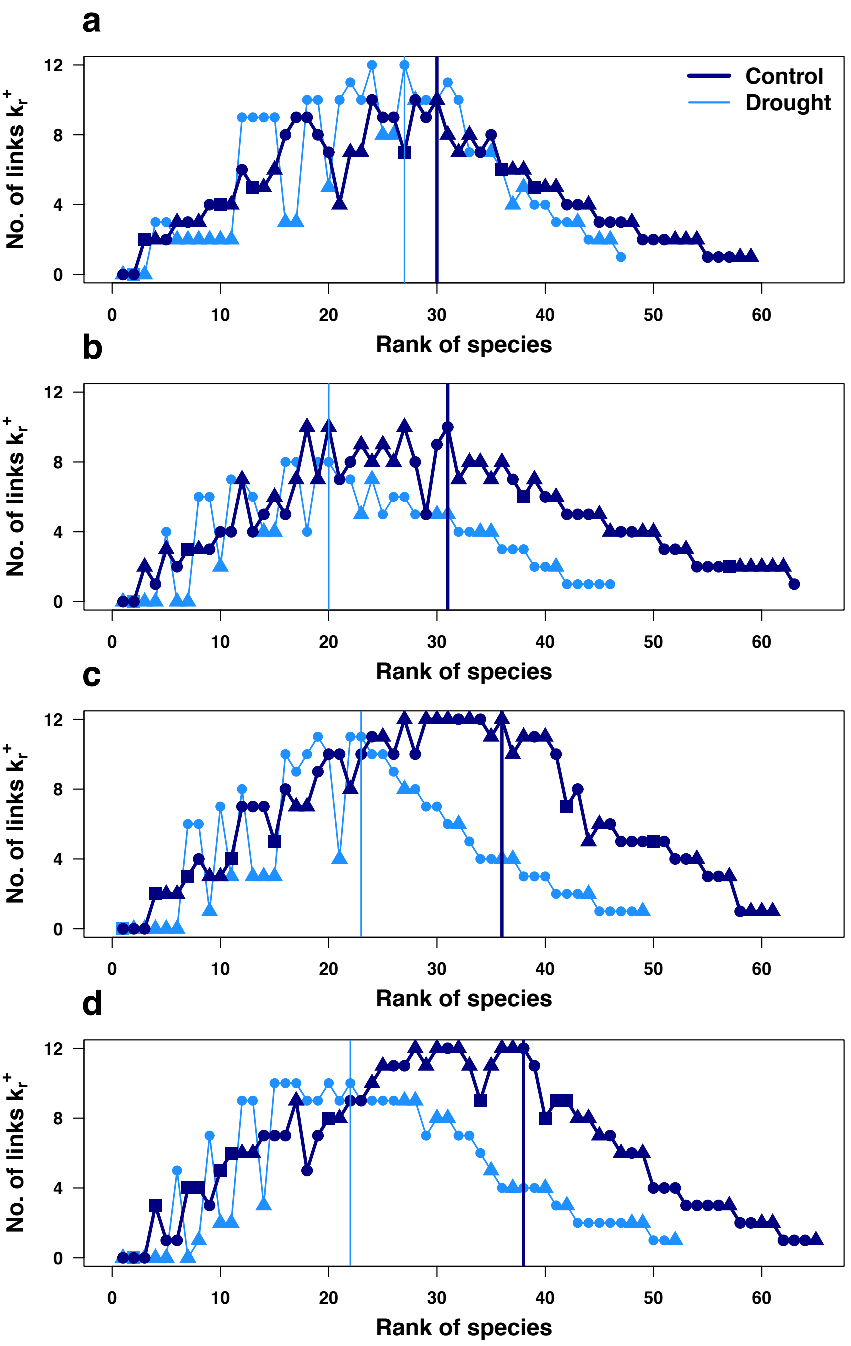


Figure 3.1: Core/periphery structure of control and drought food webs. Comparisons of four blocks of control and drought core profiles (**a-d** for web pairs 1-4 respectively). Nodes are ranked by their decreasing order of degree and plotted by the number of links with nodes of a higher rank, k_r^+. The control web (dark thick line) is plotted alongside its respective drought web (light thin line). Species were classified as basal (circles), intermediate (squares) or top (triangles). The maximum of the curve k_r^+, defines the boundary of the core for the control (dark thick line) and drought (light thin line) webs.

Table 3-C: Statistics from two independent samples t-tests. The effects of drought on the relative core (Table 3-B) and robustness (Table 3-E) were tested using one-tailed t-test on arcsine transformed data. Two-tailed t-test on arcsine transformed data were applied to examine if peripheral species are more susceptible to extinction and if more core species realigned after drought (Table 3-B). Significant effects / differences are indicated in bold (Further details described in Table 3-A).

	Drought	
	df	*p*
Relative core size	3	0.16
Robustness (random)	3	0.89
Robustness (targeted)	3	0.17
	Drought impacted substructures	
	df	*p*
More extinction from periphery	3	**0.01**
More species realigned from core	3	**0.01**

Drought caused a significant drop in degree for most of the species (shown with a positive bar length in Figure 3.2). Among the top 20 species with the largest degree reduction, only 4 of them are invertebrates whilst all the rest 16 species are the resource (i.e., *detritus* or *producer*). Among the top 5 species with the largest degree increase, 4 of them are invertebrates and only 1 of them is the producer. Results showed that resource species in general have a lower degree under drought, which can be linked to the extinction of many large predators (i.e., invertebrates); whilst the survived predators under drought have a higher degree by establishing new predation links, in order to survive under the stressed environment.

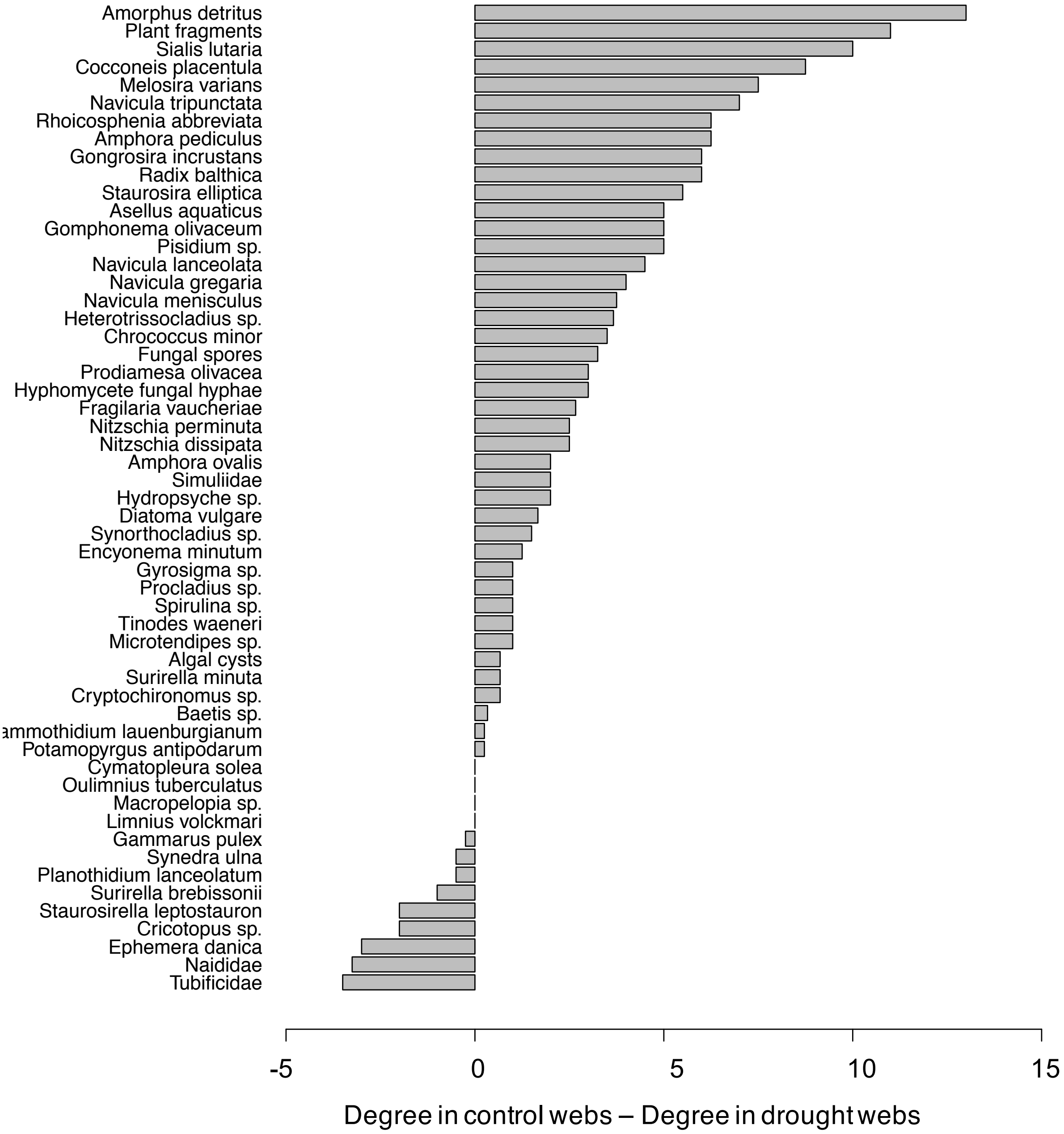


Figure 3.2: The degree change for each species that presented in both control and drought food webs. The bar is calculated as the degree of each species in control webs minus the degree in drought webs. The degree change for species presence in two or more pairs of webs were averaged.

3.3.2 Species Extinction under drought

Species extinction in substructures caused by drought were observed in all four pairs of webs. The peripheral extinctions mostly consist of invertebrate consumers located at the higher level of the food chain (Table 3-D). Specifically, for each of the invertebrate species, there is a higher chance for them to go extinct (comparing the *Frequency in periphery* with *Extinction* in Table 3-D). The absent of basal species in the drought webs was in fact caused by the extinction or diet switching of their original consumers (Table 3-D). In Figure 3.3 web pair **d**, drought caused 15 species to go extinct and 2 species to invade into the ecosystem. Similar as web pair **d**, web pair **a - c** also showed more species extinction from control webs than species invasion into drought webs. A relatively large proportion of periphery species went into extinction under drought (30% to 45%), compared to that of the core (13% to 26%, one tailed t-test on arcsine transformed proportion data, d.f.=3, p=0.01; Table 3-C).

Table 3-D: List of periphery species that went extinct under drought. Species were found to be either an *invertebrate* or a *producer.* Their frequency found in the (control) periphery and the number of times they went extinct in drought webs are listed. Species were ordered by the number of times it went extinct from the periphery

Species	Category	Frequency in periphery	Extinction
Elmis aenea	invertebrate	4	4
Ostracoda	invertebrate	3	3
Polypedilum sp.	invertebrate	3	3
Athripsodes sp.	invertebrate	3	3
Haliplus lineatocollis	invertebrate	3	3
Pentaneura sp.	invertebrate	3	3
Polycentropus flavomaculatus	invertebrate	3	3
Oulimnius tuberculatus	invertebrate	4	2
Brychius elevatus	invertebrate	2	2
Erpobdella octoculata	invertebrate	1	1
Heterotrissocladius sp.	invertebrate	1	1
Leuctra geniculata	invertebrate	1	1
Cricotopus sp.	invertebrate	2	1
Synorthocladius sp.	invertebrate	4	1
Theodoxus fluviatilis	invertebrate	1	1
Platambus maculatus	invertebrate	1	1
Sialis lutaria	invertebrate	2	1
Algal cysts	producer	4	1
Amphora ovalis	producer	4	1
Cymatopleura solea	producer	4	1
Cymbella lanceolata	producer	2	1
Diatoma vulgare	producer	3	1
Fragilaria vaucheriae	producer	4	1
Staurosirella leptostauron	producer	3	1
Surirella minuta	producer	4	1

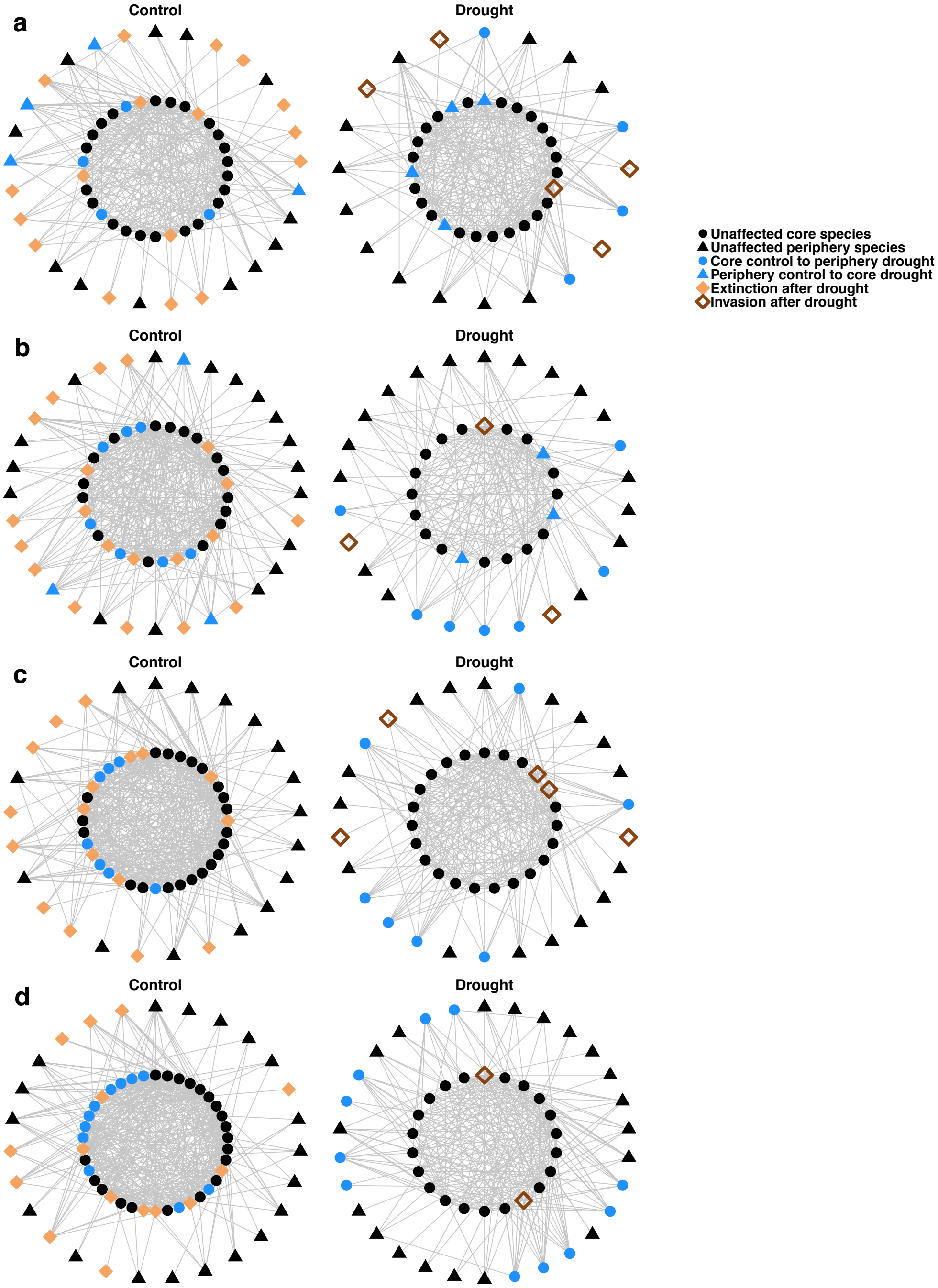


Figure 3.3: Comparisons of four blocks of control and drought food web structures (**a-d** for web pair 1-4 respectively). Drought caused species re-alignment in substructures. Core species in the inner ring are surrounded by periphery species in the outer ring. In web pair **d**, drought caused 15 species to go extinct (filled diamonds) and 11 core species to shift to the periphery (light circles).

Figure 3.4 shows four scenarios under which basal species appeared in the control webs but not in the drought webs. Those basal species all locate in periphery, which are consumed by only one or a few core species under the control condition. For example, in Figure 3.4**a**, *Cymbella lanceolata* is consumed by a single core species, and the diet of this core consumer was narrowed in the presence of drought. Similar trends are observed in Figure 3.4**b - d**.

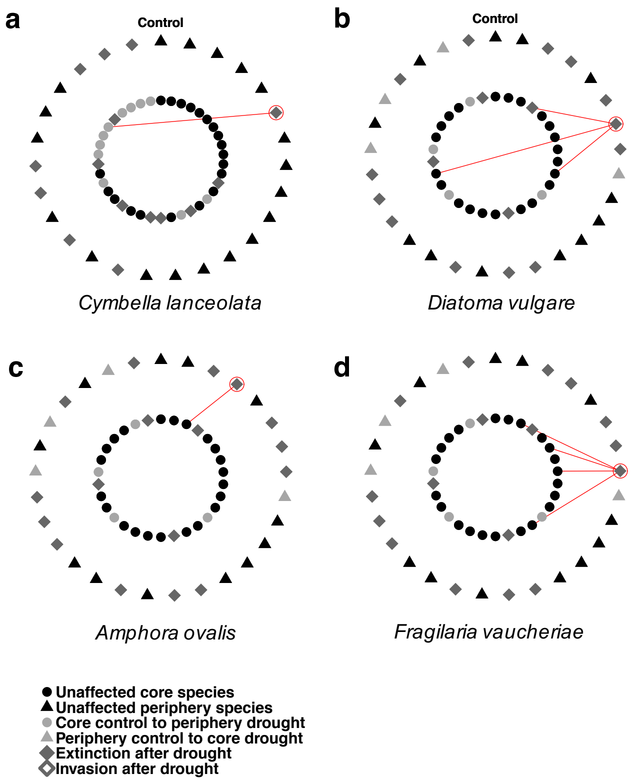


Figure 3.4: Peripheral species extinction in food webs. Species in **a - d** are all primary producers. A grey circle represents species moving from core to periphery, which indicates a reduction in its degree. A grey triangle represents species moving from periphery to core, which indicates an increase in its degree.

3.3.3 Species Movement under drought

Drought caused a larger proportion of species in the core to migrate into the periphery of the web than vice versa (one tailed t-test on arcsine transformed proportion data, d.f.=3, p=0.01, Figure 3.3 and Table 3-C). The movement between the two substructures was due to the changes in species' degree and the connectivity pattern among them. Here three species that react differently under drought were selected to demonstrate the link rewiring mechanism (Figure 3.5). The snail *Radix balthica* (Figure 3.5**a**) is tolerant of drought conditions, reflected in its diet expansion after drought. Therefore it was present in the core in both control and drought webs. The isopod *Asellus aquaticus* (Figure 3.5**b**) moved from core to periphery as degree dropped significantly after drought. The reduced number of resources likely reflects changes in the biotic habitat and encounter rate under drought. The interactions between *Asellus aquaticus* and some of its resources were destroyed by the drought stressor. For the midge *Cricotopus sp* (Figure 3.5**c**), the movement was from periphery to core as its diet diversified under drought, again, reflecting redistribution and likely altered encounter rate.

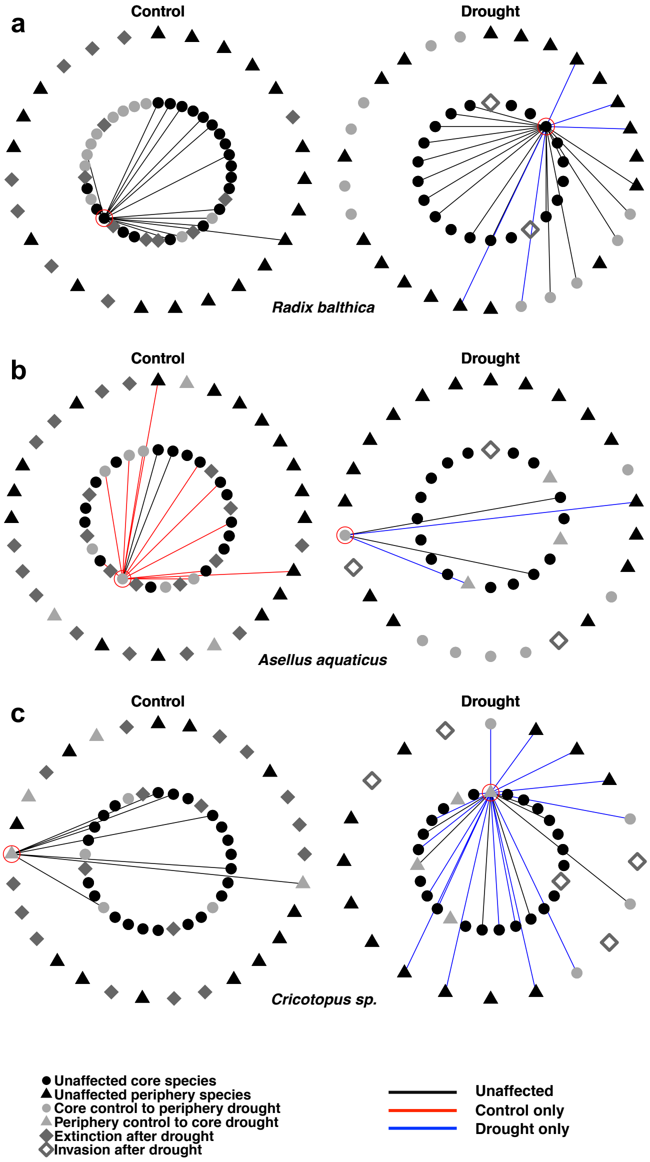
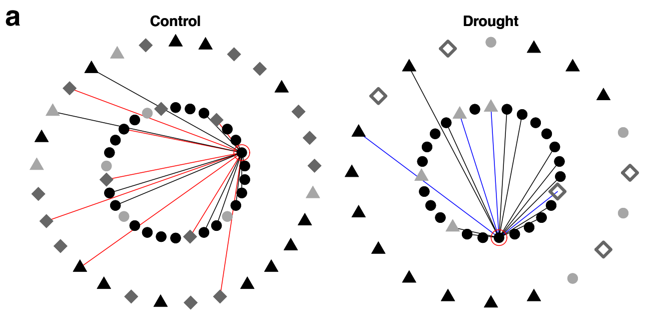
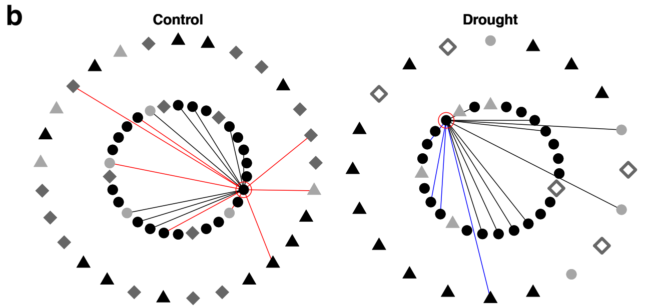


Figure 3.5: Rewiring in food webs. Core species in the inner ring are surrounded by periphery species in the outer ring. Focal species highlighted by circles. **a** - The snail *Radix balthica* is tolerant of drought conditions. **b** - The isopod *Asellus aquaticus* moved from core to periphery as degree dropped significantly after drought. **c** - The midge *Cricotopus sp* diversified its diet under drought.

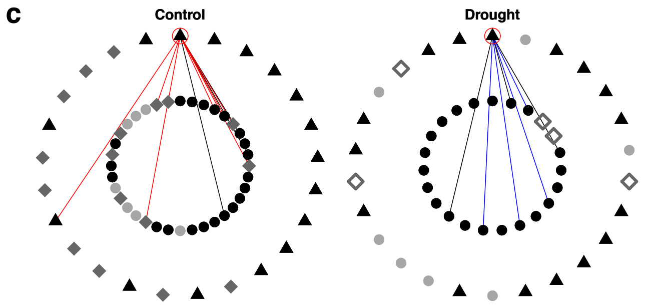
Rewiring happened not only within species that moved between substructures, but the results also shown that it was a generic responses in the majority of the species. Species that maintained their position in either core or periphery after drought were likely to experience significant link re-establishment, despite the fact that their degree did not change dramatically. In Figure 3.6, three species that stayed in the same substructure after drought were selected to demonstrate the link rewiring mechanism. *Cocconeis placentula* is a primary producer that remained in the core, despite a large amount of original interactions were lost. It kept the position within the core as new consumers switched their diet to it (Figure 3.6**a**). The similar mechanism was shared by consumers as well. Detritivore *Ephemera denica* kept its position within the core by generating new links to new prey to compensate the loss of existing interactions (Figure 3.6**b**). The decomposer *Fungal spores* (Figure 3.6**c**) stayed in the periphery under drought also with the majority of links being rewired.



Cocconeis placetula (core species)
Primary producer



Ephemera danica (core species)
detritivore



Fungal spores (peripheral species)
decomposer

● Unaffected core species
▲ Unaffected periphery species
● Core control to periphery drought
▲ Periphery control to core drought
◆ Extinction after drought
◇ Invasion after drought

Figure 3.6: Species that stayed in the same substructure after drought. Core species in the inner ring are surrounded by periphery species in the outer ring. Focal species highlighted by circles. **a** - The primary producer *Cocconeis placentula* stays in the core with large proportion of link rewiring took place. **b** - The detritivore *Ephemera danica* stays in the core with a larger proportion of peripheral link lost and a larger proportion of core links generated under drought. **c** - The decomposer *Fungal spores* stays in the periphery under drought with the majority of links being rewired.

3.3.4 Link density within the core

Drought greatly reduced the core link density (left panel of Figure 3.7 (**a-d**)). When more peripheral nodes were included, the difference in ϕ_r between control and drought webs became much smaller and finally becomes indistinguishable towards the tail of the curve in 3 out of 4 comparisons (except the first pair of food webs). All food webs showed a marked deviation in connectivity from their respective null models within their cores (Figure A1), which revealed a systematic and non-random substructure. This was shown by the significantly lower z-score at the beginning of the curve (i.e., within the core, link density inside was significantly lower than what would be expected by chance, Figure 3.7). The difference of the rich club coefficient and z-score between control and drought web were shown within the core, indicating a higher level of organisation within drought web as the response under stress. The result was consistent with previous conclusions that non-ecological networks tend to reduce their core link density under stress [170, 171].

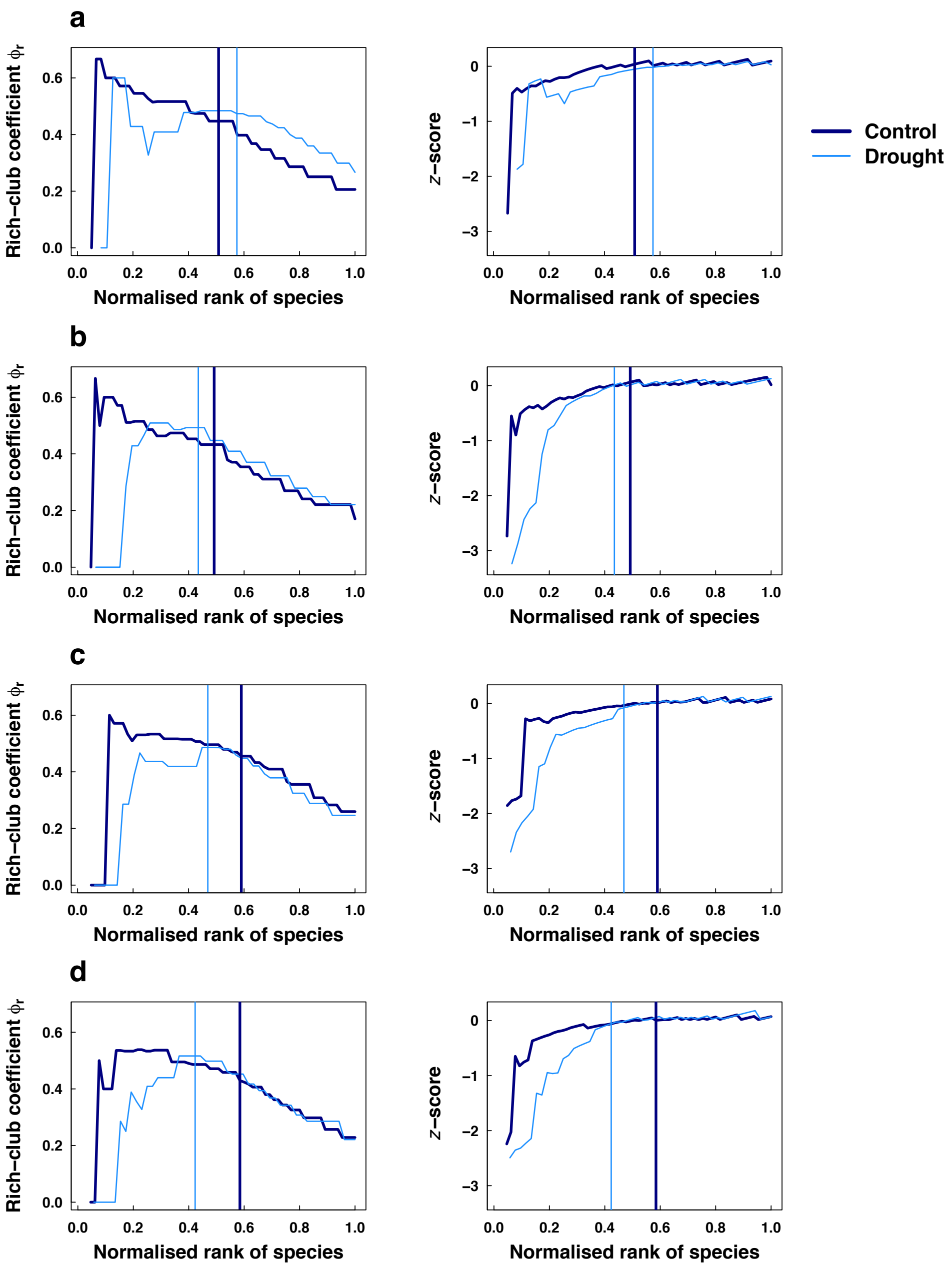


Figure 3.7: Drought reduced the link density in the core and caused further restructuring in the core. The density of connections across the network measured by the rich-club coefficient, ϕ_r, is shown for four blocks of control (dark thick line) and drought-disturbed (light thin line) mesocosms (**a-d** for web pair 1-4 respectively). Nodes were ordered by degree which were then normalised by the size of the network. Boundaries of the cores are marked by vertical lines as in Figure 3.1. Comparisons of the web pair's deviance in connection density from their respective null models and more negative z-scores indicate greater deviance from the null model.

3.3.5 Food web robustness

Food web robustness under random species removal was measured repeatedly for 100 times [169] for each food web. The average robustness, μ, and the standard deviation, σ, was shown in Table 3-E. Regardless of different experimental conditions, all the eight food webs exhibited a high average robustness. Moreover, drought did not alter the robustness significantly (t-test on arcsine transformed proportion data, d.f.=3, p=0.89; Table 3-C).

Table 3-E: Robustness of control and drought webs under simulated species removal. Proportion of species required in primary removal to generate a total of 50% species loss in each case is shown. In the case of random removal, the average robustness, μ, and the standard deviation, σ, obtained from 100 runs are shown for each empirical web.

	Random removal			
Web pair	Control		Drought	
	μ	σ	μ	σ
1	0.43	0.04	0.46	0.03
2	0.45	0.03	0.46	0.05
3	0.46	0.02	0.43	0.03
4	0.46	0.02	0.44	0.03

	Targeted removal	
Web pair	Control	Drought
1	0.25	0.32
2	0.27	0.17
3	0.31	0.20
4	0.32	0.23

When the highest degree species was removed in each iteration, food web exhibited significantly lower robustness compared to random removal, as secondary extinctions are more likely to happen after each targeted removal (Figure 3.8, Table 3-E). Similar to robustness under random removal, drought webs were as robust to species removal as control webs under targeted removal (t-test on arcsine transformed proportion data, d.f.=3, p=0.17; Table 3-C).

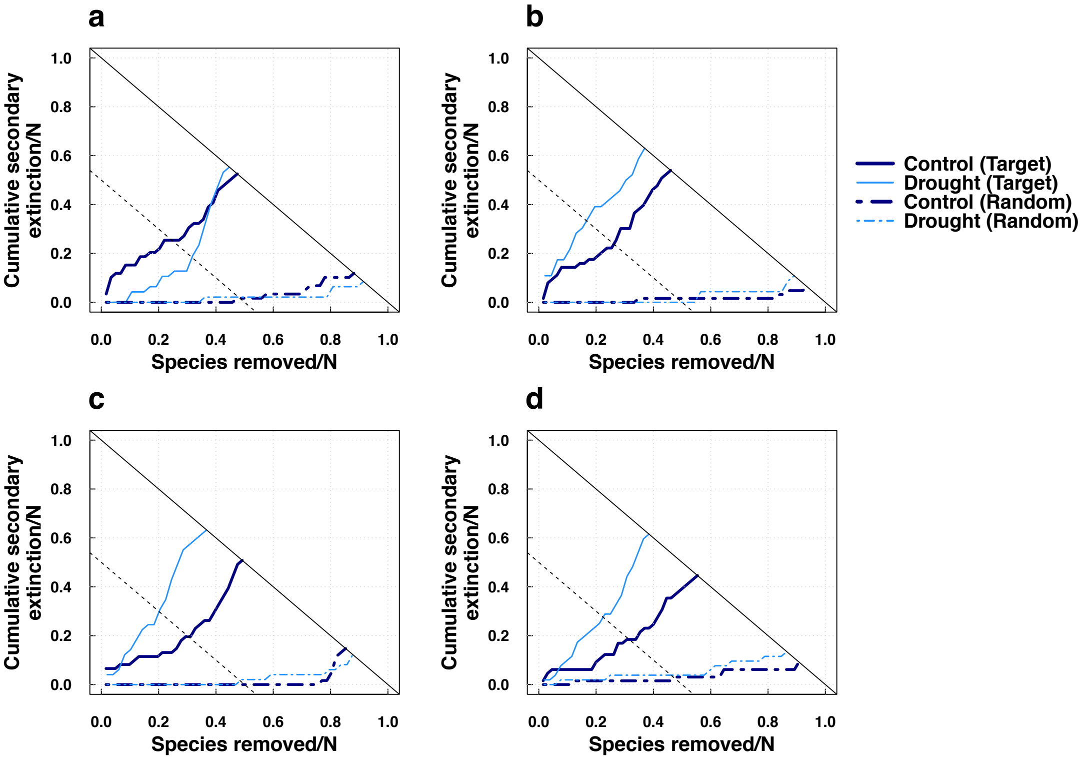


Figure 3.8: Network robustness against random and targeted species removal. Cumulative secondary extinction against simulated random species removal (dashed) and targeted generalist removal (solid) for four blocks of control (dark thick line) and drought-disturbed (light thin line) mesocosms (**a-d**). The solid diagonal line represents a total loss of 100% of species and the dashed diagonal line represents a total loss of 50% of species

The unchanged robustness (both random and targeted removal) can be explained by the preserved core/periphery structure. The relative core size, as a substructure level property, reflects how much perturbation the core can buffer through its densely intertwined pathways [170, 171]. This suggests that although the density of connections within the core was altered by drought, overall network integrity and the ability to withstand further perturbations were conserved by species re-alignment.

3.4 Summary

Drought on mesocosms streams caused significant biodiversity loss on freshwater ecosystems, among which large predators were the most vulnerable species [36, 50]. Species invasion was also observed in all the four replicates, indicating the potential devastating consequences on the long term ecosystem stability and sustainability. Despite the multiple factors that might destroy the original ecosystem, food webs within those streams were able to adapt to the changing environment and to maintain the food web structure.

Here, an important food web substructure is revealed using advanced network profiling techniques [102]. The results showed that food webs contain a richly connected core structure consisting of high degree species, which is surrounded by periphery species with a lower degree (Figure 3.3). This substructure acts as a buffer towards environmental stressors as the core contains interconnected pathways which increase the food web robustness. The discovery of this food web structure helps explain why freshwater food webs are able to persist through time when species turnover were taken place.

Specifically, the constitution of core species is more stable than periphery species as the species extinction caused by drought happened mostly to the periphery. Also, among those periphery extinctions, the invertebrates are the ones that are most vulnerable; whilst peripheral producers are relatively stable under drought. Most of the resource species experienced a reduction in degree, which is closely related to the massive extinction of the predators. The remaining predators in general showed an increase in degree, indicating that they have adapted to the stressed environment by establishing new predation links to potential prey. The conservation of robustness simulated by both random species and targeted removal indicates that network level link density plays a dominant role in determining the resilience of food webs under perturbations.

The research conducted is able to highlight that perturbations have an impact not only on individual nodes, but also on the internal substructure of the web. The substructure level changes cannot be detected using traditional analytical approaches, therefore the previously unknown compensatory dynamics are able to be explained, which is important in maintaining the food web structure and functionality.

Chapter 4

Network analysis on agroecosystems

4.1 Research background and overview

The increasing food demand across the globe [172] calls for a sustainable farming environment, to preserve farmland biodiversity and ecosystem stability. To solve this problem, food web research on agrosystem has been proposed as one of the focus areas [173]. Farming practices such as crop switching and the adoption of genetically-modified-herbicide-tolerant (GMHT) crops have been proven to affect ecosystem biodiversity [174, 175]. Existing research mainly focuses on how new farming practices would affect the abundance and diversity of arable plants and invertebrates. For example, non-pest species that are the prey of most birds and butterflies showed a continued decline in their abundance under the GMHT management [175–178]. However, existing studies only focused on the species level, for which is hard to quantify whether a change in species biomass or abundance would further affect the whole ecosystem functionalities. It is therefore not possible to measure or predict any cascade effects from species extinction or invasion.

Here, the first research goal was to construct food webs for each individual field in order to measure the impact of farming practices on both the species level and the network level properties. This research compared GMHT food webs with their corresponding conventional food webs, and this was done for all four major types of crops (i.e., beet, maize, spring oilseed rape and winter oilseed rape). Species data used to construct the network was sampled from 251 fields across the UK which covers the aforementioned four crop types [104]. Each field was split into half conventional crop varieties and half GMHT varieties of the same crop, resulting in 502 individual food webs in total.

Species level properties including the proportion of basal, intermediate and top-level species and the connectance were used to test the effect of management practices. The substructure properties such as the relative size of the richly connected core structure, the core link density and the species composition within the core and periphery respectively between management were also evaluated to gauge the variations. Finally, robustness measurement was used to examine the influence of GMHT management practices [17]. Switching crop type could significantly affect farmland biodiversity [179, 180], and differences in network properties arisen from switching crop type were used as a benchmark to gauge the variations caused by the GMHT management.

4.2 Methods

4.2.1 Data set

The Farm Scale Evaluation (FSE) is the largest research project to date to investigate the impact of GMHT on farmland biodiversity [181]. For details on the experiment design and the sampling procedure for specific crops, please refer to [174]. To summarise, the data involves 256 individual crop fields across the UK (Figure 4.1), which covers four types of crops: 64 beet (B), 57 maize (M), 65 spring oilseed rape (SR), and 65 winter oilseed rape (WR) crop fields. Within each field, two farming practices exist. Half of the area grows the conventional crop and the other half grows the GMHT variety. Strict regulations and guidelines have been applied to ensure the two halves share similar environment conditions [175, 181]. Species were sampled using both vortis suction [182] and pitfall traps [183].

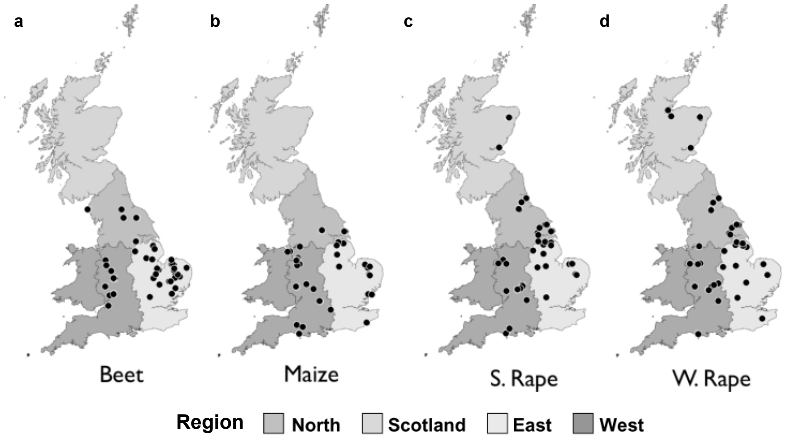


Figure 4.1: The distribution of individual FSE sites based on crop type and region within the UK [104, 177].

4.2.2 Individual and aggregated food web construction and analysis

The possible species interactions were generated based on Abductive Inductive Logic Programming (A/ILP) which examined the changes of species abundance between the half conventional and the other half GMHT field from each crop site (for details of the method, please refer to [104, 106]). The method generated all possible predations links among all the species identified (provided by Dr. David A. Bohan and Dr. Alireza Tamaddoni-Nezhad), which were the used as a reference. To generate individual food webs, the presence of a species in a given half field was determined by the species abundance and interactions between pairs of species were inferred from the reference links learnt from the A/ILP technique. The A/ILP technique is to date more accurate than other probabilistic models when generating food webs. This is because the A/ILP approach is abundance based, as opposed to most other body size based techniques [105]. For example, A/ILP can detect certain special interactions, including those that do not follow the predator-prey body size scaling rules, such as spiders and their prey.

Food webs were divided into three trophic levels (i.e., B/I/T). Species distribution was calculated as the proportion of total species contained in each level and this method was applied to all the 502 food webs. Species trophic level distribution differences between food webs were examined (i) between two treatments within the same crop type and (ii) between conventional crop types. Food web connectance as a macroscopic food web metric, was used as a 'summary' of how densely connected the food web is.

To examine species level properties from a broader prespective, for each crop type and each treatment, individual food webs were aggregated into one web, resulting in a total of eight aggregated webs. The weight of a node was calculated as how many times a species appears in all the sites, and the weight was normalised by the total number of individual fields.

4.2.3 Substructure analysis

First, for each individual food webs, core profiling method was applied [102] (Detail in Section 2.3.1). The importance of the core was examined from three perspectives. First, the relative size of the core as well as the link density within the core were used to examine whether farming practices (both crop switching and GMHT management) would affect the substructure of the food webs. Second, how many times a species was found in core or periphery were summarised and compared based on aggregated food webs, and the results could indicate if frequently appeared species are from the core. Finally, species movement was measured by comparing either among crops or between management.

4.2.4 Food web robustness

The potential effect of perturbations on network robustness was measured by simulating random removal and targeted removal, and the effect of GMHT management was benchmarked by that of different crop types [17, 82]. The random removal simulation was run for 100 times and results were averaged. For targeted removal, when a node was removed from a food web, the degrees among the rest of the nodes were re-calculated [71, 130, 184, 185].

4.2.5 Statistical analysis on network properties

The Bray-Curtis index [186] (Section 2.5.4), b, was applied to quantify the compositional similarity between any two aggregated webs with reference to the total counts of each species. b was also used to quantify the compositional similarity between core and periphery of each aggregated webs. When $b = 0$, it indicates the highest similarity between the two webs such that each species has the identical appearance frequency, and $b = 1$ as the most dissimilar, such that any species that appears in one aggregated web is absent in the other one.

The GMHT management practices differ between crop types, that is, the influence of the GMHT management on one type of crop is not the same with GMHT management on another type of crop. Therefore, a nested Type I ANOVA, with crop management nested within crop types, was used to examine if there is any significant difference on species composition properties caused by management within the same crops. To account for pseudo-replication, an error structure with each half field nested within each site was used. To test the effects of different crop types on food web properties, a type II one-way ANOVA was used on conventionally managed food webs only. Both models were applied to food web properties (proportion of B/I/T, connectance, relative core size, core link density, and food web robustness) and significant results were followed by Fisher's LSD *post hoc* test to identify the corresponding factors.

4.3 Impact of agricultural practice on species level properties

In each aggregated web (Figure 4.2), the variation was found across all the trophic levels. Bray-Curtis index [186], b, showed that species were less similar across crop types than that when comparing conventional with its GMHT counterpart of the same crop type (Figure 4.3). The comparison showed that small variations were observed within each site (i.e., between the conventional and GMHT crops), and cross-site dissimilarities were more significant. This implies that variations in species composition observed previously [177] between conventional and GMHT webs were localised effects and did not transform into systematic changes in the wider food web, whereas crop type had far more profound effects.

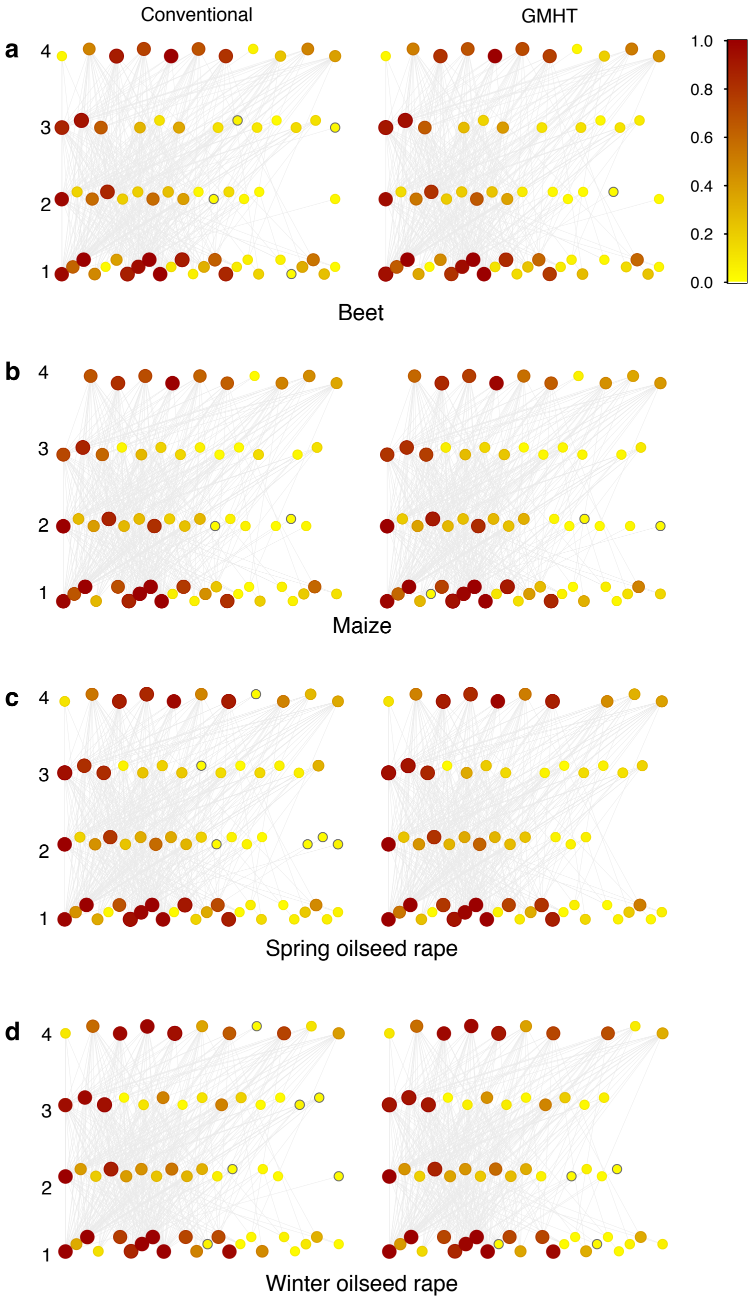


Figure 4.2: Pairwise conventional and GMHT webs (**a** - beet; **b** - maize; **c** - spring oilseed rape; **d** - winter oilseed rape) with the same species placement between each conventional and GMHT pair. Node size and colour denote the proportion of times a species was found in the given crop variety across all the sites. Nodes bounded by a dark edge are unique to their respective webs.

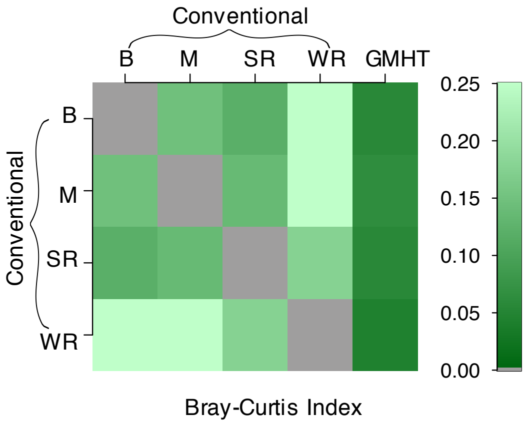


Figure 4.3: Comparisons of species dissimilarity between crop types and management using the Bray-Curtis dissimilarity index. Colour denotes the degree of dissimilarity with b = 0 as the most similar and b = 1 as the most dissimilar.

4.4 Impact of agricultural practice on food web network properties

4.4.1 Food web properties

In this section, individual food webs (i.e., 502) were examined. Food webs of a given crop variety (e.g., Conventional or GMHT) differed greatly in size across sites, varying between 12 to 38 species. Yet food webs of the two varieties of a given crop type were highly correlated in size (Figure 4.4). Linear regression was used to examine the correlations in web size between individual conventional webs and their GMHT counterparts, and they were shown to be significantly correlated ($0.48 < R^2 < 0.61$).

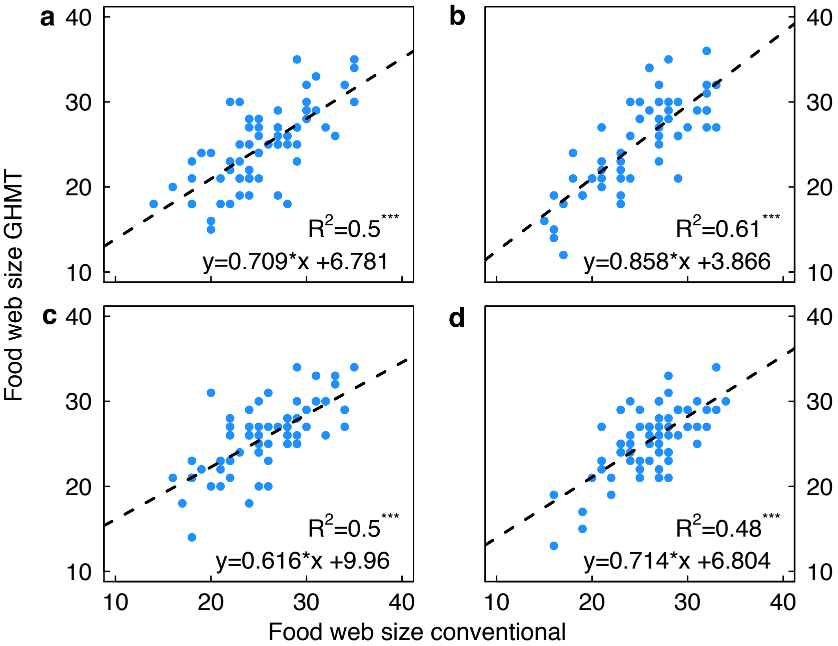


Figure 4.4: The size of conventional food webs was plotted against the size of the corresponding GMHT half-fields (**a** - beet; **b** - maize; **c** - spring oilseed rape; **d** - winter oilseed rape). The dashed line denotes the linear regression, with its function and R-square shown. *** Regression significant at $P < 0.001$.

The proportion of basal species and intermediate species were unaltered by management, but the proportion of basal species (P = 0.030; Table 4-A) varied significantly among crop types. *Post hoc* tests showed the significance was specifically caused by the conventional beet has a much larger proportion of basal species than the conventional winter oilseed rape (Table 4-A). However, the proportion of top predators varied significantly between management (P = 0.020). This result was due to the significant difference between beet and maize crops (GMHT beet > conventional beet, conventional maize > GMHT maize, Table 4-A). Food web connectance was greater under GMHT, and *post hoc* tests showed that this was caused by the connectance of GMHT beet webs were larger than that of conventional beet webs (P = 0.023).

Table 4-A: ANOVA on the effects of management and crop type on food web structural properties. Nested ANOVA on the effects of management within each crop type. One-way ANOVA on the effects of crop type among conventional crops. Significant ANOVA results were further analysed using Fisher's LSD *post hoc* tests. %B, the proportions of basal species; %I, the proportion of intermediate species; %T, the proportion of top predators; C, connectance.

	Management					Crop type				
	df	*SS*	*MS*	$F_{4,247}$	*P*	*df*	*SS*	*MS*	$F_{3,247}$	*P*
%B	4	0.002	< 0.001	0.12	0.975	3	0.033	0.011	3.02	**0.030**[‡]
%I	4	0.010	0.002	0.68	0.608	3	0.024	0.008	2.39	**0.069**[*§]
%T	4	0.014	0.004	1.97	**0.020**[‖]	3	0.005	0.002	1.01	0.390
C	4	0.006	0.002	2.85	**0.023**[†]	3	0.004	0.001	2.38	**0.070**[*§]

[*] Marginally significant ($0.05 < P < 0.09$)

[‖] Conventional beet > GMTH beet, Conventional maize > GMHT maize.

[†] GMHT beet > conventional beet

[‡] Conventional beet > conventional winter oilseed rape

[§] Conventional winter oilseed rape > conventional beet

4.4.2 The core and its link density

For each individual food web, the core/periphery structure was identified and network properties related to this substructure were examined. Here, the core was seen as a *rich-club* which consists of high degree nodes that play a dominating role from both topological view and functional view. The method used to profile the core firstly ranked the species based on degree and examined the interconnectivity among high degree nodes sequentially. The boundary of the core was detected when the core link density reaches its peak and decreases afterwards. The details of how to profile a core is described in Methods, section 2.3.1. This *two-class partition* would always identify a core and a periphery in non-random networks. As a consequence, all the 502 food webs possessed a core surrounded by loosely connected peripheral species [102] (Figure 4.5), with the conventional webs and their GMHT counterparts sharing a large proportion of the core and periphery species. Relatively large cores were observed across all the food webs compared to non-ecological networks, accounting for an average of 65-71% of the total species (Figure 4.5).

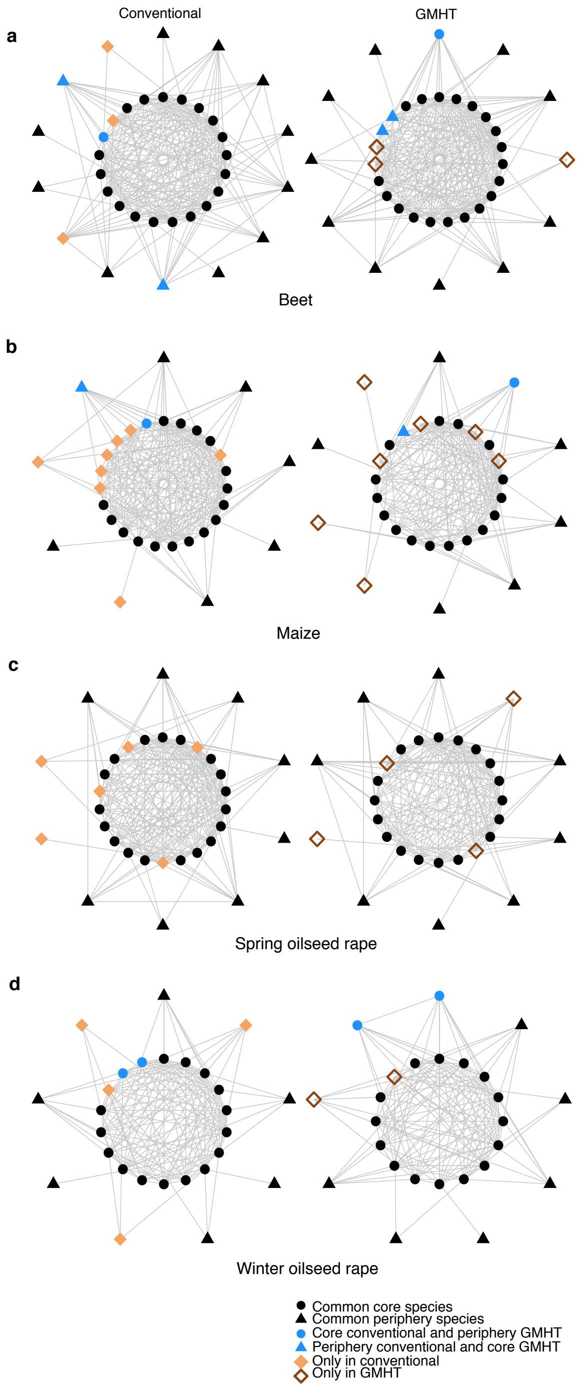


Figure 4.5: Core structure was highlighted as the inner circle, surrounded by peripheral nodes. One of the selected site was plotted across four crop types (**a -** beet; **b -** maize; **c -** spring oilseed rape; **d -** winter oilseed rape). For each node, the color and shape indicate species movement patterns between conventional webs and their GMHT counterparts.

The link density (the rich-club coefficient ϕ_r, see Section 2.3.2) within the core of both conventional and GMHT food webs share similar trends (Figure 4.6). The larger ϕ_r observed within the core indicates that higher link redundancy confers stability upon the food webs in the face of external perturbations [170]. The similar trend of ϕ_r also indicates the connectivity among high degree species contributes to the formation of similar core size across all the individual webs.

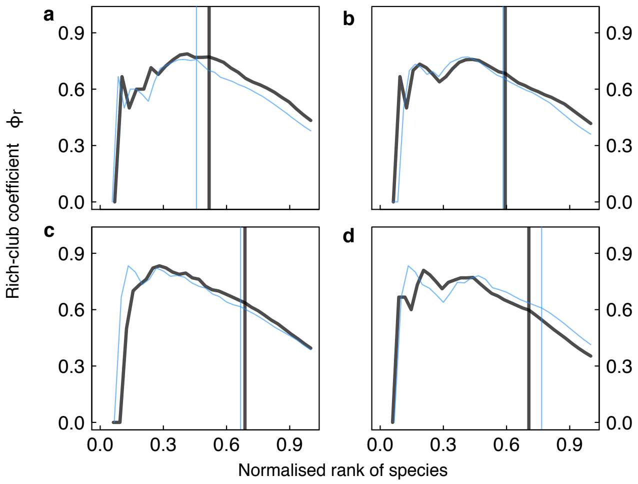


Figure 4.6: The rich-club coefficient (ϕ_r) between conventional (black thick line) and GMHT (light thin line) was compared for one of the selected site across four crop types (**a -** beet; **b -** maize; **c -** spring oilseed rape; **d -** winter oilseed rape). Nodes were ordered by the decreasing order of degree and normalised in order to make webs with different sizes comparable. The vertical lines represent the boundary of the core correspondingly.

ANOVA analysis showed no significant differences in the relative core size when examining the effect of conventional and GMHT management ($P = 0.521$; Table. 4-B), however, when the effect of different crop types was considered, marked difference was observed ($P = 0.002$; Table. 4-B). The *post hoc* test (Fisher's LSD) pointed out the significance was caused by the maize crop whereby the related food webs have larger cores than those of both beet and winter oilseed rape. The link density within the core also showed a similar trend with the relative core size, where no significance was found when comparing between management (Figure 4.6; $P = 0.547$; Table. 4-B) whilst marked difference was found among crop types ($P < 0.001$; Table. 4-B). In this case, the *post* *hoc* test showed the significance was caused by the winter oilseed rape, with higher core link density than maize and spring oilseed rape. As both the relative core size and core link density are indicators of network robustness [56, 170] which reflects their resilience against perturbations, a difference in either properties would potentially result in marked difference in their ecosystem functionalities.

Table 4-B: Analysis of variance on the effects of management and crop type on food web substructural properties and robustness. Nested ANOVA on the effects of management within each crop type. One-way ANOVA on the effects of crop type among conventional crops. The relative core size, core link density (the ϕ_r), and robustness via random removal and targeted removal are shown. Significant results were further analysed using Fisher's LSD *post hoc* tests.

	Management				
	df	*SS*	*MS*	$F_{4,494}$	*P*
Relative core size	4	0.032	0.008	0.81	0.521
Core link density (the ϕ_r)	4	0.010	0.002	0.77	0.547
Robustness via random removal	4	0.005	< 0.001	0.48	0.750
Robustness via targeted removal	4	0.007	0.002	0.62	0.649

	Crop type				
	df	*SS*	*MS*	$F_{3,247}$	*P*
Relative core size	3	0.143	0.048	4.87	**0.002** †
Core link density (the ϕ_r)	3	0.064	0.021	6.80	<**0.001**‡
Robustness via random removal	3	0.002	0.001	2.54	**0.057***§
Robustness via targeted removal	3	0.024	0.008	2.93	**0.034**‖

* Marginally significant ($0.05 < P < 0.09$)

† Maize > beet; maize > winter oilseed rape, all conventional.

‡ Winter oilseed rape > maize; winter oilseed rape > spring oilseed rape, all conventional.

§ Conventional spring oilseed rape > conventional beet.

‖ Conventional winter oilseed rape > conventional beet.

4.4.3 Substructure species composition and turnover

Taxonomic composition in core and periphery was compared via: 1) the conventional and its GMHT counterpart and 2) different crop types. The common core (or periphery) species for each crop type represents the species that appears in both the conventional and GMHT counterpart (Table 4-C). Most core species that appear in conventional webs also appear in the GMHT core (average number was around 14, Table 4-C, **I**, row 1 & 2). However, the number of common species in periphery of conventional and GMHT webs was much smaller (average number was around 5), as very few species experienced large degree changes between conventional and GMHT web (Table 4-C, **I**, row 3 & 4). Apart from species movement, there also exist cases when species from the core (or periphery) only appears in either the conventional or GMHT food web. The variations in species composition are slightly higher in periphery, and this is consistent for both treatments (Table 4-C, **II**, row 6 & 8). This indicates the species composition is much more stable in core than in periphery in all the cases.

Table 4-C: Comparisons on the species composition in the core and periphery. **I:** For a given crop, the number of common core or periphery species was compared between conventional and its GMHT counterpart. A small number of species were found in the core of the conventional webs but in the periphery of the GMHT counterparts, and vice versa. **II:** The number of species that was unique to either conventional or GMHT web was similar, and the similarity hold when these unique species were further split into core and periphery substructure. Regardless of conventional or GMHT web, the number of unique species in core was in general larger than that in periphery.

I:	B	M	SR	WR
Common core species	13.81 ± 2.86	14.46 ± 3.81	14.55 ± 3.16	14.48 ± 2.95
Common periphery species	5.41 ± 2.51	4.12 ± 2.18	4.92 ± 2.02	5.54 ± 1.99
Conventional core & GMHT periphery	0.86 ± 1.17	1.23 ± 1.64	1.29 ± 1.78	0.85 ± 1.20
Conventional periphery & GMHT core	0.95 ± 1.37	0.88 ± 1.23	0.98 ± 1.17	1.29 ± 1.73
II:				
Core species in conventional web only	1.80 ± 1.51	1.95 ± 1.51	1.69 ± 1.41	1.57 ± 1.37
Periphery species in conventional web only	2.69 ± 1.77	2.19 ± 1.61	2.38 ± 1.81	2.25 ± 1.50
Core species in GMHT web only	1.69 ± 1.25	2.00 ± 1.27	2.05 ± 1.45	1.42 ± 1.17
Periphery species in GMHT web only	2.14 ± 1.68	2.47 ± 1.90	2.06 ± 1.50	1.77 ± 1.30

Species compositional similarity in the core and periphery were further examined using the Bray-Curtis index b. Within the core, crop type has greater impact on the species composition, as the smallest b all result from management comparison (Figure 4.7**a**). The largest b appears when the core of conventional beet was compared to the conventional winter oilseed rape. Similar trend of dissimilarity was shown when the periphery is compared (Figure 4.7**b**). However, the overall b of periphery was consistently higher than that when comparing among the core, indicating the rigidity of the core in maintaining its components and the plasticity of species movement in the periphery.

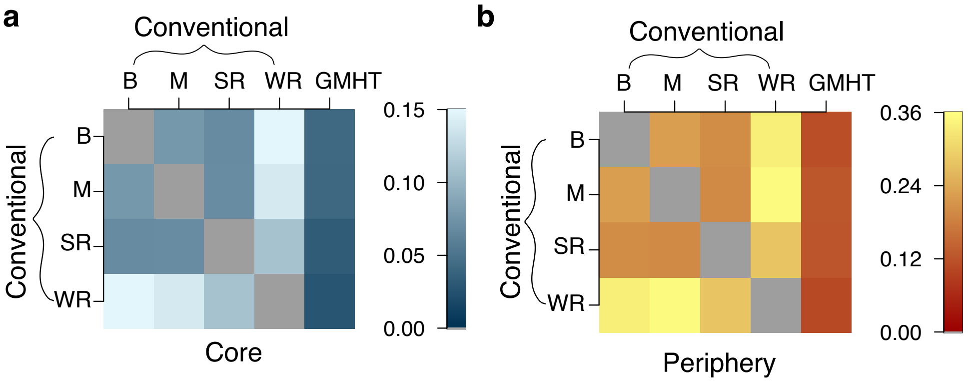


Figure 4.7: Similarity in the species composition among individual webs grouped by crop types and management, quantified using the Bray-Curtis dissimilarity index, was overall higher in **a** - core than **b** - periphery. Within the core and periphery substructure respectively, the similarity is higher when comparing conventional with GMHT webs than comparing across crop types.

4.4.4 Recurrent core substructures

There exists small unions of substructures, consisting the same species, that appear at a significantly high frequency across all the individual webs within the same crop category. The present of this recurrent substructure may have implications on what dominant roles would a network exhibit [145, 187].

The recurrent substructure consists of species that always appear in the core of both conventional and GMHT treatment (Figure 4.8). The union of this substructure is the key for driving community and ecosystem properties given their central location (i.e.,

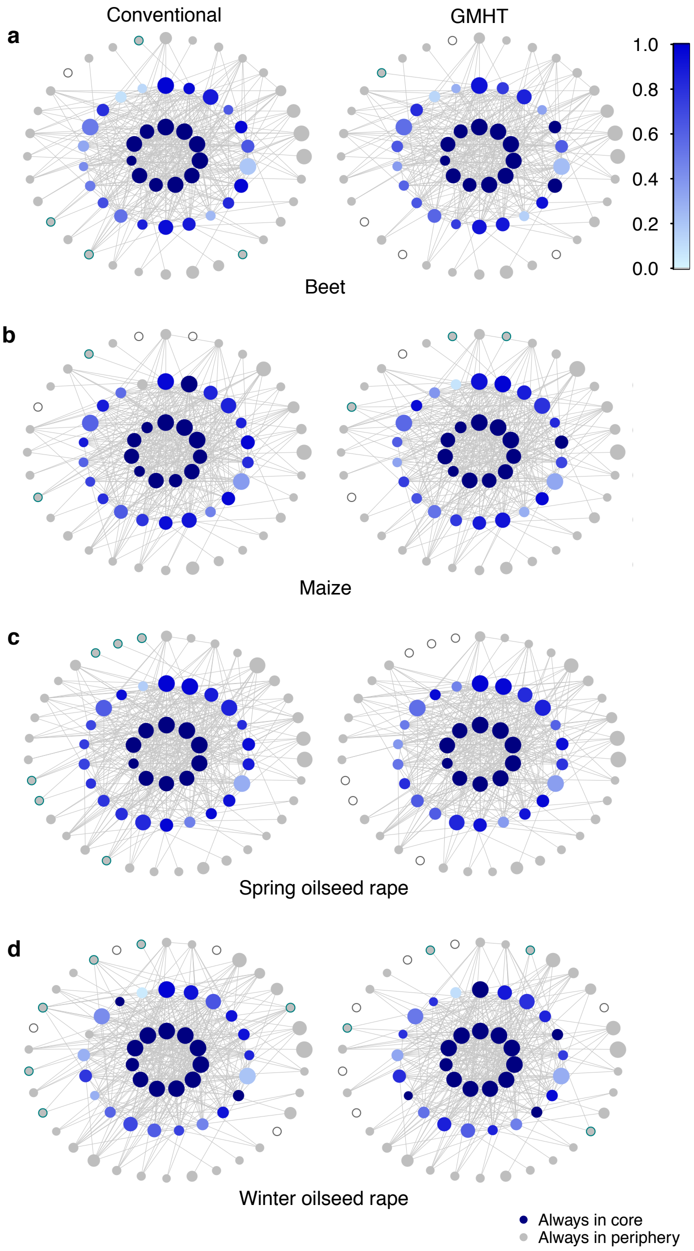


Figure 4.8: Pairwise conventional and GMHT webs with the same species placement between each conventional and GMHT pair (**a** - beet; **b** - maize; **c** - spring oilseed rape; **d** - winter oilseed rape). Node size denotes the proportion of times a species was found in the given crop variety across all the sites. Colour denotes the gradient of core presence (white indicates absence). Species that were always found in the core in both conventional and GMHT are in the inner ring, and similarly, species that were consistently found in the periphery in both conventional and GMHT are in the outer ring. The rest of the species are in the middle ring. Nodes bounded by an edge denote absent species (unfilled) and species that were unique to their respective web (filled).

core nodes with high degree) in the network and their prevalent appearance in all the crops [188]. The middle circle consists of species that not always appear in the core. The outer circle consists of species that never appears in the core. The larger the node, the more times it was found in all the individual webs, the darker the node is, the higher probability that the node was found in the core. The location of each species is fixed between conventional and GMHT counterpart, therefore, it clearly shows which species is absent from the current web while present in the other.

4.5 Impact of agricultural practices on food web robustness

The architecture of food webs governs their stability and underpins their response to perturbations [36]. To evaluate whether different farming practices would cause a difference in the fragility of ecosystems, food web robustness was measured for each individual food webs, and results were compared. ANOVA analysis showed that management had no effect on either robustness under random removal or targeted removal which can be explained by their homologous network structures (Table. 4-B), while crop type has led to marked differences ($P = 0.057$, $P = 0.034$, Table. 4-B). For the random removal, the standard deviation σ of each 100 simulations within each crop type and crop variety is around 0.076. All the results in Table. 4-B indicate clearly the effect from crop switching is significantly stronger than that from different crop managements. The network-level responses of GMHT are similar to that of conventional webs (Figure 4.9), comparing to the distinct differences among crop types. Each metric studied in this chapter is summarised in Figure 4.9, which is normalised by its overall range across all webs. The effects of crop type can be visualised by comparing results horizontally. The more towards the center, the smaller the metric is compared to that of other groups.

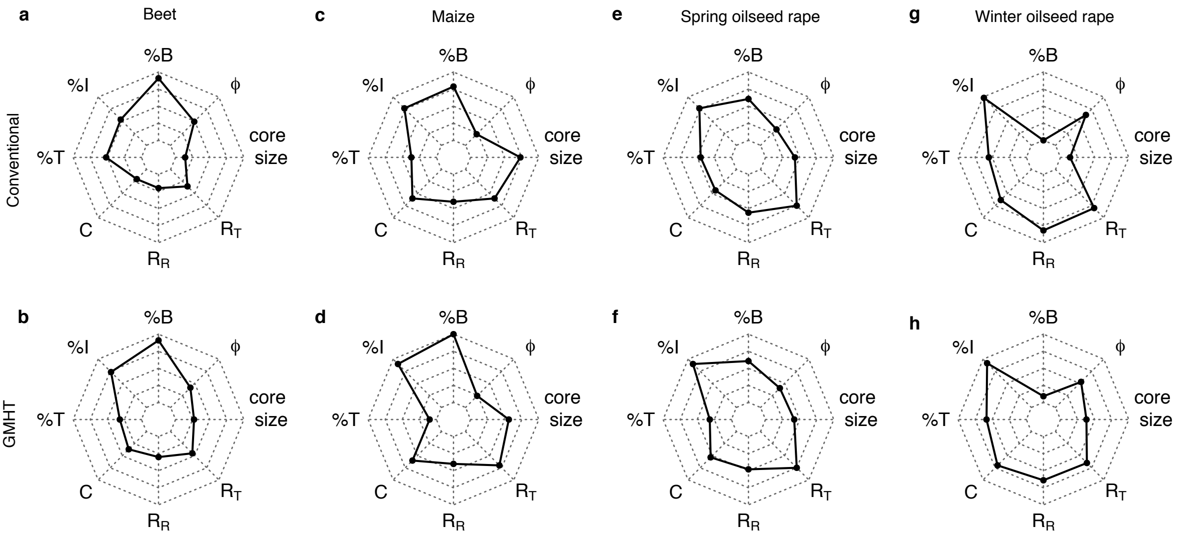


Figure 4.9: Pairwise comparisons between Conventional and GMHT management varieties (**a,b** - beet; **c,d** - maize; **e,f** - spring oilseed rape; **g,h** - winter oilseed rape). %B, the proportions of basal species; %I, the proportion of intermediate species; %T, the proportion of top predators; C, connectance; core size; ϕ, core link density; R_{R}, robustness via random removal; R_{T}, robustness via targeted removal. Each metric is normalised by its overall range across all webs. The effects of crop type can be visualised by comparing results horizontally.

4.6 Summary

Along with climate change, land-use practices have been claimed to have the most profound effect on the biodiversity loss of terrestrial ecosystems [189]. Indeed, land-use results in severe damage to the habitat of species. Crop type is a known cause of biodiversity change in farmland [180], and this was used as a benchmark to gauge the relative effects of management regimes including the conventional treatment and the GMHT treatment.

One of the widely believed ecological risks caused by GMHT crops, which was studied here, is the impact on farmland biodiversity loss [190]. The underlying reasons that biodiversity disturbances observed previously in GMHT crops were largely due to the focus on the species level. Here, the importance of the core was reflected from the following three perspectives: First, the core tends to resist change, while the periphery is active during adaptation, reflected in the common higher proportion of core species between the two counterparts. Aggregated food webs also showed that low degree peripheral species share a larger dissimilarity. Core species were more commonly found than peripheral species and formed highly recurrent substructures involving fixed species members (Figure 4.8). Second, the relative size of the core does not differ significantly between treatment, indicating a similar level of core redundancy. This contributes to the same level of tolerance under perturbations. Finally, the core connectivity (the rich-club coefficient) between conventional and its GMHT counterparts vary similarly within the same crop, solidifying the fact that the unchanged core size is not an artefact of overlap farmland species, but instead is due to the similar connectivity pattern among high degree species.

The network based analysis presented here suggests that the changes in species abundance presented in previous works should not have been interpreted as a potentially important risk to the agricultural ecosystem. New farming practices on GMHT crops exert a relatively small effect on both the food web substructure (relative core size, core link density) and robustness (under both targeted and random removal) compared with the switching of crop types which is already widely accepted in farming (Figure 4.9). The study provides insights for applying novel network measurements and techniques to better assess food webs under different agroecological environments.

Chapter 5

Core redundancy governs food web robustness

5.1 Research background and overview

Complex network research across many disciplines has revealed the importance of the network topology [54, 58, 66, 95, 116]. In particular, network topology is strongly linked to its robustness in the face of perturbations, such as removal of the central hub in scale-free networks can lead to large-scale cascading effects, whilst the removal of the other lower degree nodes can only cause minor local effects which do not propagate through the whole network [71, 95, 191].

In ecology, robustness has been used as a metric to quantify the level of perturbations a network can withstand before it collapses [17]. The most widely used network level metric which links to robustness is connectance [38, 42, 77]. In general, an increase in the food web connectance is caused by the increase of the average number of prey per consumer, which results in a lower probability that the survival of the consumer is dependent on a specific prey [17, 43, 79]. However, the predictive power of connectance is limited, because the connectivity of individual species could be different even when different networks have the same connectance [42]. For example, a food web contains only a few basal species providing resources to support all higher level species, the extinction of which would destroy the network [37, 133]. Different environmental conditions (e.g., pH level, drought) can lead to distinct trophic-level structures among food webs, yet the difference cannot be indistinguishable through connectance [11, 15, 36, 81, 192].

In food webs, redundancy is defined as one or more resource species being able to fulfil the role (i.e., energy transfer) of an existing resource if it was removed [12, 82, 121, 125, 126], which minimises the potential secondary extinction of a predator when the system is perturbed, such as environmental warming [122], acidification [81, 123] and drought [4, 124]. Despite the fact that food web redundancy is important to maintain robustness, no topological based measurement for food web redundancy has been defined so far. Therefore the amount of redundancy within a food web and its effects towards maintaining robustness are still unknown. Here, the notion of a spanning tree [12] was applied to characterise the level of redundancy in food webs. The spanning tree structure was also used to describe the distribution of the energy or information flow, and the proportion of redundant links in the core/periphery substructures.

5.2 Methods

5.2.1 Data set

This set of data (containing 53 individual food webs) were sampled from different ecosystems, including classic food webs which are extensively studied (e.g., Benguela, Broadstone stream [193], Skipwith Pond, Ythan Estuary, and Tuesday Lake sampled in 1984 and 1986 respectively), mesocosms food webs treated under different environment conditions (e.g., 8 Mill Stream food webs under drought and control conditions), Kennet food web experienced the environmental stress caused by pesticide, UK stream food webs exposed to different degrees of acidification (e.g., 20 highly resolved food webs over a wide pH range [81, 108]), and 18 US riverine food webs sampled within the ecoregion near Ohio states, resulting in 53 food webs in total (Table A1). The ecosystems cover streams, rivers, lakes, ponds, and marine. The 20 pH food webs were provided by Dr. K. Layer. The US riverine food webs were provided by Dr. C. Mulder. The Millstream food webs were provided by Dr. M. E. Ledger. The classic food webs were from the Cheddar Package [107]. Food webs in this chapter are considered as un-weighted graphs. And unlike previous chapters, for the generation of spanning trees, edges in the graphs are directed.

5.2.2 Constructing a spanning tree

In graph theory, a spanning tree is a sub-network of the original network, which connects all the nodes with the minimum number of links. When applied to food webs, the spanning tree is a directed and acyclic graph originating from the bottom-level of the food web and reaches all the top-level species with every two nodes being connected by only one link [146]. If there are two or more species providing resource to a predator, one of the links is randomly selected to be part of the tree structure and the rest to be redundant (Algorithm 1). The links in the spanning tree structure reflect the plausible paths through which energy flow is delivered to each species [12]. Different trees can be constructed from the same network; therefore 1000 trees were generated for each empirical food web and results were averaged.

With this type of re-construction, the complexity caused by 'loops' is reduced, and the focus is then shifted towards the the minimal number of links that are needed to maintain the food web structure while making sure all the species are connected. The resultant tree can be seen as a typical topological structure of networks that has *zero* redundancy, since the functional property of redundant links is to provide alternative paths between species. In most of the real networks, predators tend to have more than one prey within the ecosystem, thus they should contain certain level of redundancy.

```
Input   : Directed binary food web $G$ consists of $N$ species and $E$ links.
Output: One possible construction of spanning tree $G_\mathrm{s}$ consists of $N_\mathrm{s}$ species
          ($N_\mathrm{s} = N + 1$) and $E_\mathrm{s}$ links. The set of links is represented as $\mathcal{E}\mathrm{s}$.
initialisation:
$N_\mathrm{s} = 1$ ($N_\mathrm{root}$ is added);
$\mathcal{E}\mathrm{s} \leftarrow [N_\mathrm{b} \leftarrow N_\mathrm{root}]$;
$N_\mathrm{s} = 1 + N_\mathrm{b}$;
while $N_s < N + 1$ do
    Randomly select a predator $n_\mathrm{non.b}$ from $N \cap N_\mathrm{s}$ consuming an arbitrary species
      ($n_a$) in current set of $N_\mathrm{s}$;
    Select the link $E_\mathrm{s}$ connecting $n_\mathrm{non.b}$ and $n_a$ ($E_\mathrm{s} = [n_\mathrm{non.b} \leftarrow n_a]$);
    $\mathcal{E}\mathrm{s} \leftarrow E_\mathrm{s}$;
    $N_\mathrm{s} = N_\mathrm{s} + 1$;
end
```

Algorithm 1: Spanning tree generation algorithm [12].

A synthetic network is used to demonstrate how a spanning tree structure is constructed. The synthetic network $G_a = (N, E)$ (where $N = 10$, and $E = 15$) was plotted in trophic diagram in Figure 5.1**a**, and the trophic level of each species was shown on the left axis. Figure 5.1**b** shows one possible layout of the derived spanning tree $G_\mathrm{s} = (N_\mathrm{s}, E_\mathrm{s})$, which was constructed based on Algorithm 1. The construction of the spanning tree began with one *root* node manually added to the network. Directed links were generated from the *root* node pointing to the basal level species (n_1, in-degree: $k_\mathrm{in} = 0$), which indicates the energy flow from the environment (*root*) to the producers within this ecosystem. Next, a non-basal species was randomly selected from the predators, and one of the predation link was randomly selected and added to set $\{E_\mathrm{s}\}$. For example, for n_5 that feeds upon n_1, n_2, n_3 and n_4, the link between n_5 and n_1 were randomly selected to be included in the tree structure. This procedure was repeated until all the species were linked to one of their prey in G_s. Empirical links $e_{2-5}, e_{2-3}, e_{2-4}, e_{3-5}$

and e_{4-5} were considered to be topologically redundant thus removed. Finally, the set $\{N_\mathrm{s}\}$ contains all the species that were shown in the original graph G. If G contains N nodes, $\{N_\mathrm{s}\} = N + 1 = 11$ (all the original nodes plus the *root* node) and 10 links. For each food web, 1000 spanning tree structures were generated in order to achieve statistical significance [194]. Figure 5.1**a** can be re-organised into Figure 5.1**c**, which highlights the core structure (the inner ring) surrounded by peripheral nodes.

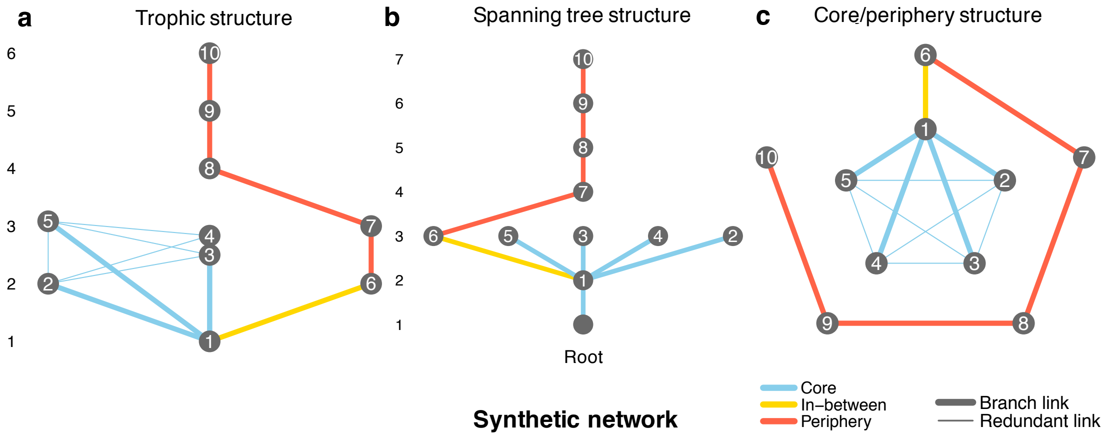


Figure 5.1: **a -** Trophic diagram layout of the synthetic network. **b -** A spanning tree structure of the network in **a** with only *branch links* included. **c -** The core/periphery layout containing a 5-node-clique as the core ($1 \sim 5$, each node is linked with the rest of the nodes) and a 5-node-periphery ($6 \sim 10$). Blue links represent core links, yellow for in-between links that connect nodes between core and periphery, and red for periphery links. For **a, c**, thick links are *branch links* while thin links are redundant links. For **b**, only branch links are included in the spanning tree.

5.2.3 Quantifying core and periphery redundancy

In this thesis, the number of redundant links in a network is defined as:

$$E' = E - (E_\mathrm{s} - E_\mathrm{root}) \tag{5.1}$$

where E_root is the number of links in the spanning tree structure that connects root node with basal species. Since the root links do not represent real predation relationship between species but only used to represent the energy flow from the outside environment to basal species, they are not considered in the calculation of food web redundancy. Thus, $E - (E_{\mathrm{s}} - E_{\mathrm{root}})$ calculates all the redundant links in the original food web.

Based on Eq. 5.1, redundant links can be further grouped into E'_{core} as the number of redundant links within the core of a food web, E'_{between} as the number of redundant links connecting core and periphery, and $E'_{\mathrm{periphery}}$ as the number of redundant links only in the periphery. Each in-between link connects a core species with a periphery species, thus the weight of in-between links was further divided such that half of the link weight is assigned to the core and half to the periphery [58].

Core redundancy:

R_{core} is calculated as the number of redundant links within the core (E'_{core}) and half of the weight of the in-between links (E'_{between}), and normalised by the total number of redundant links (E').

$$R_{\mathrm{core}} = \frac{E'_{\mathrm{core}} + 0.5E'_{\mathrm{between}}}{E'} \tag{5.2}$$

Periphery redundancy:

$R_{\mathrm{periphery}}$ is calculated as the number of redundant links within the periphery ($E'_{\mathrm{periphery}}$) and half of the weight of the in-between links (E'_{between}), and normalised by the total number of redundant links (E').

$$R_{\mathrm{periphery}} = \frac{E'_{\mathrm{periphery}} + 0.5E'_{\mathrm{between}}}{E'} \tag{5.3}$$

R_{core} and $R_{\mathrm{periphery}}$ add up to 1. Redundancy $R \in [0, 1)$ because in theory it is possible to have a food web with no redundant link, therefore, $R = 0$, whilst all the food webs contain one or more *branch links*, therefore, $R \neq 1$.

The core links and periphery links, and the links between the two substructure can be clearly seen from Figure 5.1**c**. There are in total 6 redundant links E' ($E'_{\mathrm{core}} = 6$, $E'_{\mathrm{between}} = 0$,and $E'_{\mathrm{periphery}} = 0$) resulting in the core redundancy $R_{\mathrm{core}} = (E'_{\mathrm{core}} +$

$0.5E'_{\text{between}})/E' = 1$ and periphery redundancy $R_{\text{periphery}} = (E'_{\text{periphery}}+0.5E'_{\text{between}})/E' = 0$.

For each food web, the final core and periphery redundancy was calculated by taking an average on the 1000 trees.

5.2.4 Food web robustness

Robustness was again measured in two ways: targeted removal of high degree nodes and random removal [4, 16, 17, 42, 78]. Specifically, targeted removal of a single high degree species each iteration provides estimation of how much perturbation a food web can withstand, which also indicates the importance of redundant links among high degree nodes. The targeted removal of the core substructure as a whole can show if the peripheral on its own can be a fully functional substructure (i.e., contains energy flow from prey to predators). The results were compared with randomly removing the same number of nodes. Similarly, by performing targeted removal of the periphery as a whole, it can show whether the core substructure can fully function on its own. Again, the results were compared against randomly removing the same number of peripheral nodes.

5.3 Redundancy in substructures

5.3.1 Spanning tree in empirical food webs

An empirical food web (Benguela) was selected from the 53 food webs to illustrated the calculation of core and periphery redundancy (Figure 5.2**a**), and its re-constructed spanning tree is shown in Figure 5.2**b**. The Benguela food web contains 29 nodes and 196 links, within which the core consists of 23 nodes and 168 links and the periphery consists of 6 nodes and 6 links. The rest 18 in-between links connect core nodes with peripheral nodes.

Based on the spanning tree structure in Figure 5.2**b** which contains 27 branch links that belong to the original food web and 2 branch links that connect root node with the basal species, $R_{\text{core}} = (E'_{\text{core}} + 0.5E'_{\text{between}})/E' = (151 + 0.5 \times 15)/169 = 0.93$ and $R_{\text{periphery}} = (E'_{\text{periphery}} + 0.5E'_{\text{between}})/E' = (3 + 0.5 \times 15)/169 = 0.07$ (Eq. 5.2 and Eq. 5.3).

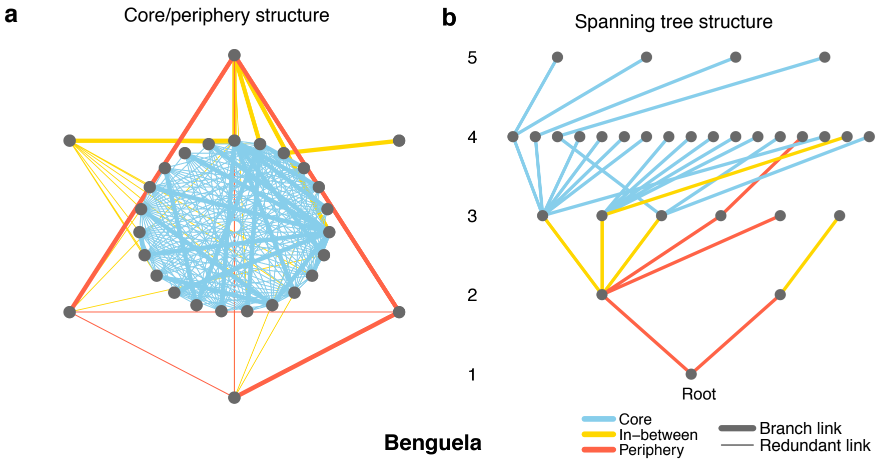


Figure 5.2: **a -** Core/periphery structure of a real food web (Benguela). **b -** The spanning tree structure of Benguela. Blue links represent core links, yellow for in-between links that connect nodes between core and periphery, and red for periphery links. Thick links are *branch links* while thin links are redundant links. Only branch links are included in the spanning tree.

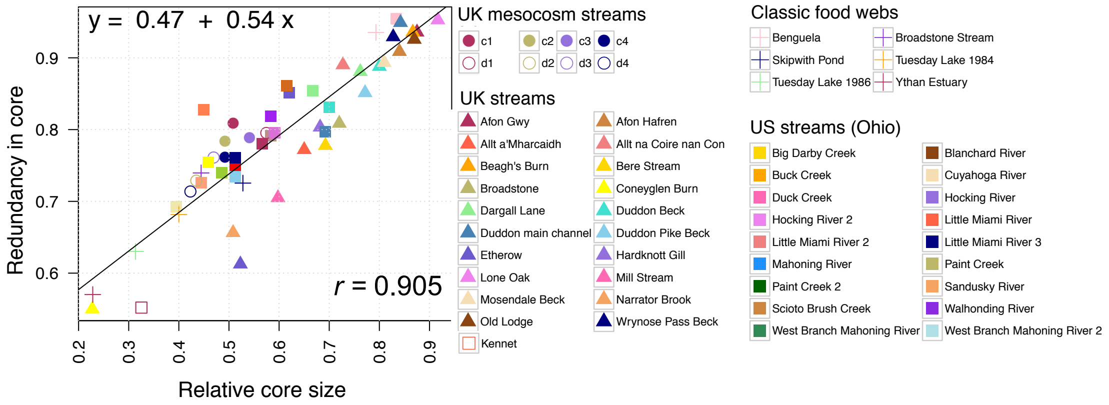


Figure 5.3: Correlation between the relative core size and core redundancy. Each food web is shown with a unique symbol, and they are grouped based on different studies. The six classic food webs are represented in cross, the eight mill stream food webs are represented using circle, the twenty UK freshwater food webs under different pH levels are represented using triangles, and the eighteen US riverine food webs are represented using squares. The Kennet food web is represented in hollow square.

A core/periphery structure was detected in all the 53 empirical food webs, with the relative core size shown in Table A1 [54, 99]. Within the 53 food webs, the core size distributes evenly from 0.2 to 0.9 in Figure 5.3, with the relative size of the core of each food web highly correlated with the core redundancy (Figure 5.3, $y = 0.47+0.54x, r = 0.905$). The larger the core is, the higher proportion of redundancy the core would contain. Although the core and periphery is similar in size, and the relative size of the core is only slightly larger than that of periphery (Figure 5.4**a**), the substructure redundancy showed marked differences (Figure 5.4**b**). The highest/lowest core redundancy are the Cuyahoga River ($R_{\text{core}} = 0.955$), the Coneyglen Burn and Kennet food web ($R_{\text{core}} = 0.55$, Figure 5.5), showing the lowest core redundancy is even greater than the highest periphery redundancy.

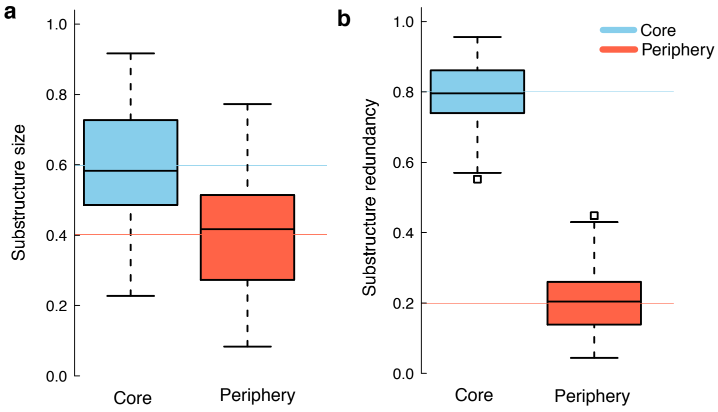


Figure 5.4: **a -** Distribution of relative core and periphery size. **b -** Distribution of food web redundancy in the core and periphery. Within the food web substructure: core and periphery, redundancy in the core (R_{core}, blue) and periphery ($R_{\text{periphery}}$, red) are summarised across all the 53 food webs (Details in Figure 5.5). The two horizontal line indicates the average of substructure redundancy respectively.

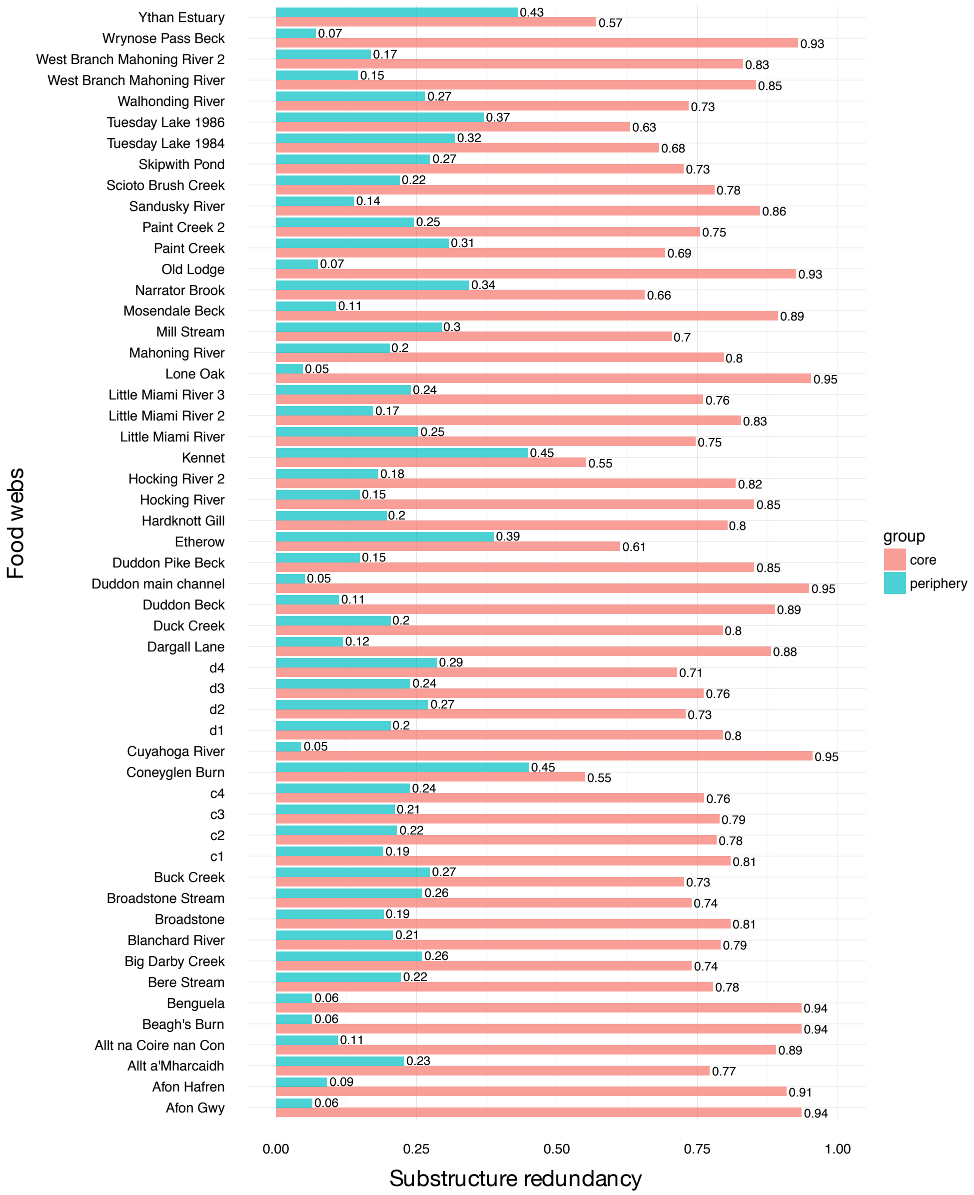


Figure 5.5: For each food webs, core redundancy (R_{core}), calculated as the proportion of redundant links in core divided by the total number of redundant links (red) is disproportionally larger than periphery redundancy ($R_{\text{periphery}}$). A summary on the distribution of core/periphery redundancy of the listed 53 food webs are shown as boxplot in Figure 5.4**a**.

5.3.2 Stream food webs under different pH levels

The relative core size of all the 20 pH streams ranges from 0.227 to 0.917, which covers 86% of the maximum range of core size among 53 food webs, whilst the connectance of all the 20 pH streams ranges from 0.114 to 0.275, which only covers 48% of the connectance range (from 0.026 to 0.037, Table A1). Previous studies found no relationship between connectance in stream food webs with different pH levels [81], here, it has been demonstrated that the relative core size is a more effective indicator for describing network properties.

Figure 5.6 shows that stream food webs with the lowest pH tend to have relatively large cores therefore higher redundancy within the cores. Here, 6 individual pH food webs (three with lowest pH and three with highest pH) were selected to show that food webs with lower pH tend to have a larger core. Food webs with lowest pH are: the Old Lodge ($cs = 0.87, R_{\text{core}} = 0.926, pH = 5.0$), the Lone Oak ($cs = 0.917, R_{\text{core}} = 0.953, pH = 5.2$), and the Beagh's Burn ($cs = 0.867, R_{\text{core}} = 0.936, pH = 5.3$, Figure 5.6**a-c**). While streams that have the highest pH tend to have smaller cores and lower redundancy, for example, the Mill stream ($cs = 0.6, R_{\text{core}} = 0.705, pH = 8.4$), the Bere Stream ($cs = 0.7, R_{\text{core}} = 0.778, pH = 7.5$), and the Hardknott Gill ($cs = 0.682, R_{\text{core}} = 0.803, pH = 6.0$), and Narrator Brook ($cs = 0.508, R_{\text{core}} = 0.656, pH = 6.0$, Figure 5.6**d-f**). Food webs with highest pH also have more peripheral nodes and more in-between links than those with lower pH.

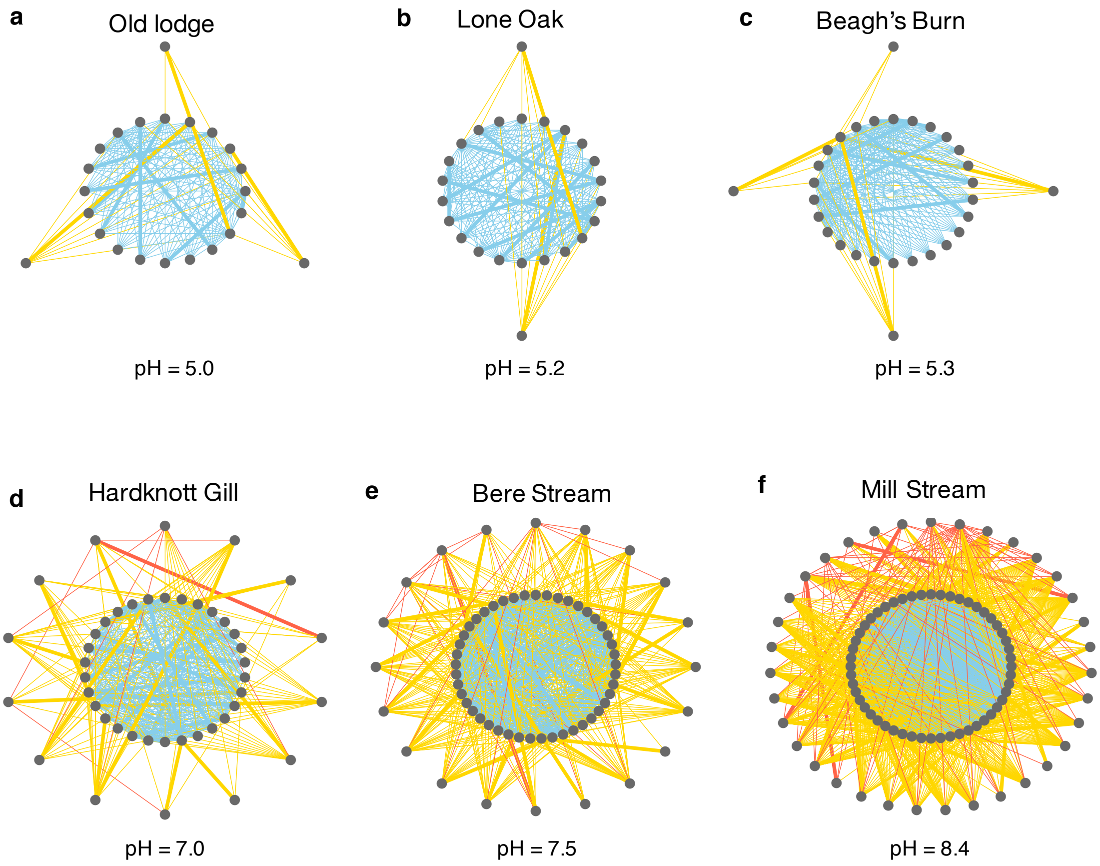


Figure 5.6: **a-c** Three food webs with lowest pH compared against another three food webs **d-f** with highest pH. The relative size of the core is larger among low pH food webs, and the in-between links are more prevalent among high pH food webs. The thick lines represent the *branch links*, while the thin lines are the redundant links.

5.4 The importance of core substructure to robustness

5.4.1 Correlation between core size and robustness

In this section, the worst-case scenario of food webs undergoing perturbations was simulated by removing species with highest degree and examined the consequences. This simulation was repeated for each of the food web, and it was found that robustness in general increases with rising relative core size. The linear fit between robustness and the relative core size is $y = 0.06 + 0.38$, with $r^2 = 0.5$, and correlation $r = 0.7044$

(Figure 5.7**a**). The correlation with robustness using relative core size is noticeably higher than that of connectace ($y = 0.15 + 0.68x$, with $r^2 = 0.28$, and correlation $r = 0.526$, Figure 5.7**b**). The stronger correlation between robustness and the relative core size indicates the effectiveness of using substructures to characterise food web functional properties.

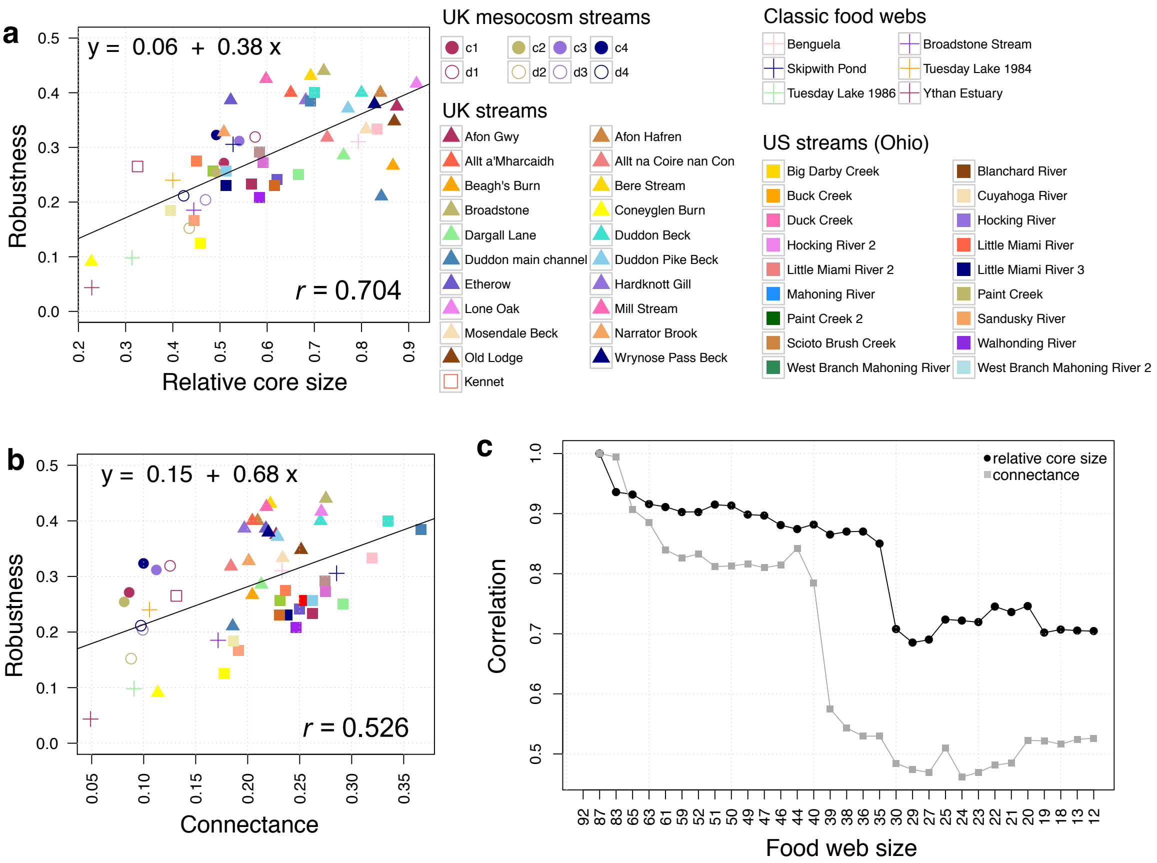


Figure 5.7: Correlation between robustness and **a -** connectance **b -** relative core size for all the 53 food webs. Both the linear regression and Pearson correlation analysis were performed and results are shown in each subfigure. Relative core size shows a stronger correlation with robustness than connectance. **c -** The change in the correlation between robustness and the two network properties (relative core size and connectance) when smaller webs were sequentially included in the analysis. The drop in the correlation is much more susceptible for connectance (grey, square) than the relative core size (black, circle) when lower threshold of the web size drops.

The 53 food webs studied range from large marine food webs (92 species - Ythan Estuary) to relatively small stream webs (12 species - Hocking river, Table A1). For such a diverse range of food webs, it is important to further examine the influence of food web size to the correlation strength between robustness and the two network properties (relative core size and connectance). The inclusion of smaller food webs reduces the correlation with robustness for both core size and connectance (Figure 5.7**c**), and this is a result of lower taxonomic resolution in smaller webs, as they are comprised of aggregated nodes rather than individual species [44, 115]. Nevertheless, the relative core size achieves almost a consistently high correlation across all web sites and is therefore a more effective measure to gauge robustness than connectance.

Table A1 provides a summary of network properties of all 53 food webs, ranked by increasing order of network size (N). Five network properties include 1) N, 2) connectance, 3) the relative core size, 4) core redundancy, and 5) robustness. All the food webs are connected graphs, with no isolated species included in the analysis.

5.4.2 Robustness under random removal

Previous studies on measuring robustness showed that food webs in general experience significantly fewer secondary extinctions from random removal when compared to targeted removal [17, 83, 133, 135]. Robustness under random removal was measured 100 times for each food web, with both the mean and the distribution (boxplot) of R_{random} shown in Figure 5.8. The results were ranked based on the relative core size, where Lone Oak food web ranked first (largest core presented), and so forth.

Most food webs exhibited an extremely robust behaviour under random perturbations. This is reflected in the fact that a few or no secondary extinction took place (shown as the median of R_{random} ranges between 0.45 and 0.5 in Figure 5.8). In addition, R_{random} was unaffected by the difference in the relative core size, exhibiting a consistently high average value and small variances across different permutations of random removal for each food web. Only three food webs (i.e., Benguela, Skipwith Pond,

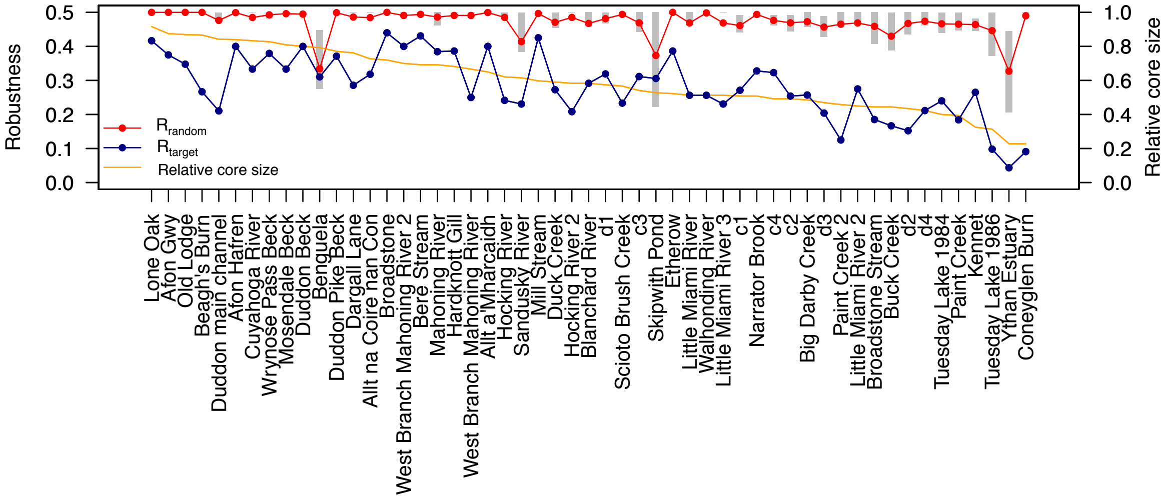


Figure 5.8: The targeted removal (blue, R_{target}, from the highest degree node) causes more secondary extinction than its equivalent random removal (red, R_{random}). The gray bar attached to each red node represents the upper and lower boundary of the random removal. Food webs were ranked based on the decreasing order of the relative core size.

and Ythan Estuary) have $R_{\text{random}} < 0.4$, since they have the fewest number of basal level species. Indeed, possessing relatively few number of basal level species increases the chance of their removal and result in cascading effects towards higher level species. The results indicate most of the food webs are highly tolerant to random perturbations, regardless of their web size and their robustness under targeted removal.

5.4.3 Substructure removal

For each food web, the consequence of substructure removal was recorded as the number of secondary extinction (N_{SE}). The proportion of secondary extinction ($\widetilde{N}_{\text{SE}} = N_{\text{SE}}/N$) shows the impact that removal of species within a specific substructure would cast. $\widetilde{N}_{\text{SE}} = 0$ means that the removal of species did not cause any further extinction or isolation of species. The larger $\widetilde{N}_{\text{SE}}$ indicates the removal of species would cause a larger effect. Results were compared against randomly removing the same number of core and periphery species, and the randomisation was performed 100 times for each food web. Figure 5.9 shows that the secondary extinction was significantly high when the core species were removed. The significance was demonstrated by comparing the result with randomly removing the same number of species in order to eliminate the bias caused by the variances in the number of core species. The equivalent random removal was performed for both core and periphery.

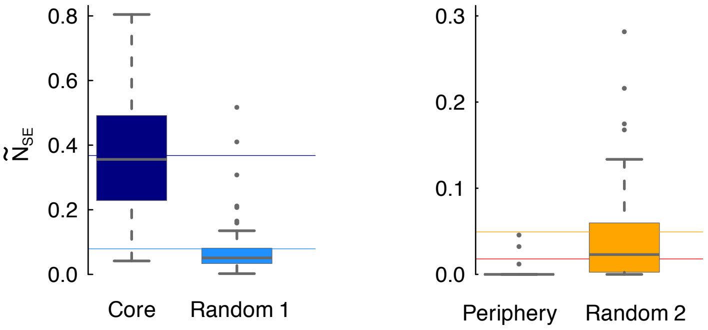


Figure 5.9: The removal of core (dark blue) causes more secondary extinction (higher $\widetilde{N}_{\mathrm{SE}}$) than its equivalent random removal (light blue). The median is represented as the black line within each boxplot, and the average of each group is shown as the horizontal line in corresponding colours.

The removal of all the periphery species caused almost no secondary extinction (median at 0, average around 0), while the peripheral equivalent random removal showed further collapse to the network. Both random removals generate similar level of secondary extinction, which indicates the relative size of core or periphery does not have an significant effect to the results (Figure 5.9). For $\widetilde{N}_{\mathrm{SE}}$ of individual food webs, please refer to Figure 5.10.

The food web core contains most of the interactions and is responsible for energy exchange with periphery, therefore the removal of the core would cause devastating effects to the network. The removal of the core disconnects the whole network, resulting in that the remaining periphery nodes on its own cannot sustain a network structure. On the other hand, the periphery has links mostly to the core rather than linking with

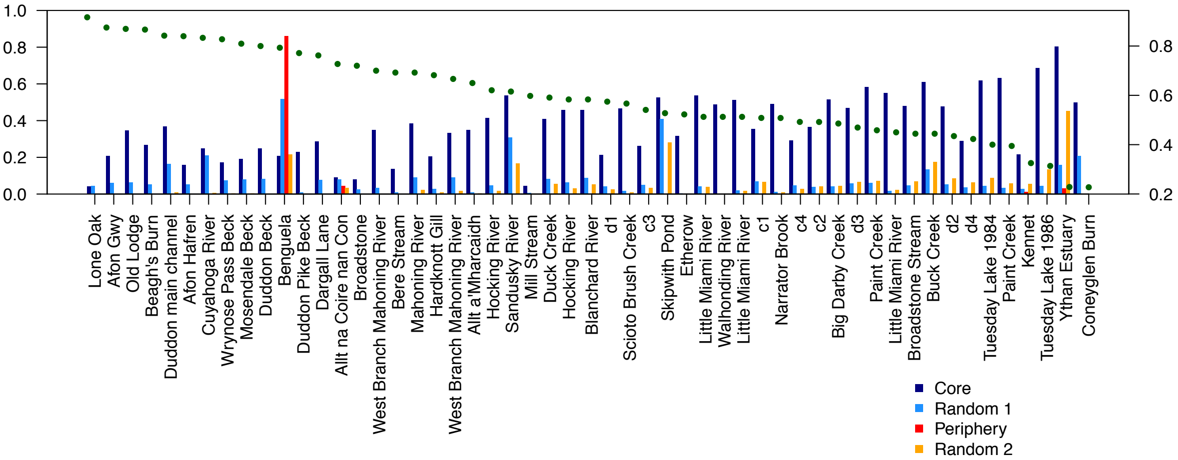


Figure 5.10: For each of the food web, the re-calculated robustness R_o is overall the lowest when performing the core removal (dark blue), whilst R_o being the highest when performing the periphery removal (red). Results were benchmarked by the equivalent random removal for both cases (light blue and orange). Food webs were ordered by decreasing order of relative core size (green dots), the left most food web has the largest core whist the right most one has the smallest core.

each other, therefore, the removal of the periphery would not affect the connectivity of the rest of the network. As such, the highly redundant and densely interconnected core can buffer the perturbations.

5.5 Summary

The importance of having a large relative core size to maintain system robustness has been shown in various types of food webs [55, 90, 170] (Lu, 2016). Here, the redundancy which is the key to buffer perturbations was found to be largely preserved within the core, and the size of the core is shown to reflect the level of redundancy in a food web. The importance of core redundancy was further explored by generating a relationship between the relative core size and robustness. Compared with connectance, the relative core size showed a much stronger correlation with robustness, and the strength of this correlation is not affected by the size of the food webs.

The dominant effect of the core was proved to be significant regardless the relative size of the core. The core is not only densely connected within the substructure, but also responsible for connecting all the periphery nodes together. The removal of the core as a whole can destroy the whole network, whilst the removal of the periphery almost results in no secondary extinction. The perturbation originated from periphery is mitigated when it reaches the core, which prevents further propagation to the rest of the network.

Chapter 6

Centrality and food web robustness

6.1 Research background and overview

Species and the interactions between them form food webs, which always contain one or more key species that have either higher abundance or unique trophic position that is of importance to maintain the stability and sustainability of an ecosystem [21, 22, 108, 195]. The extinction of a key species can seriously affect the survival of other species, which might ultimately lead to the collapse of the ecosystem [17, 22, 84, 195]. A way to quantify node importance using a network approach is via centrality [19, 20, 85, 196]. There are many ways to define whether a species is of importance to the whole food web, for example, high degree species are those that have many interactions with other species, therefore, the removal of high degree species is able to cause more secondary extinctions compared to removing low degree ones [17, 80]. High degree species are also found with higher probability to be presented in more fragments (e.g., geometry niches) within a certain geographical area [147]. On the other hand, high betweenness species are those that lie on the majority of the shortest paths and they are often considered as bridges between different subcommunities within a single food web [61, 154]. Girvan & Newman proved that the removal of the highest betweenness species in Chesapeak Bay food web resulted in the splitting of pelagic and benthic sub-communities [61]. High closeness species are those that have a short distance to all the other species in terms of the number of hops between them. In host-parasite networks, the parasites with the highest closeness centrality possess the best location to accumulate resources [154]. Eigenvector centrality measures whether a node is connected to a more influential node, therefore can be seen as a global indices compared to the local degree centrality [87, 158].

Identifying keystones for conservation in ecology has been focused on rare species [197, 198]. There have been numerous studies trying to find important species in food webs in a topological way [17, 19, 61, 83, 144, 196]. Degree centrality has been the first and most widely accepted metric to identify key species [17, 148]. Later studies expand this idea by also considering the neighborhood degree when finding the keystone species [20, 85]. However, there is no linkage between the key species identified using these methods and species ecological properties.

53 food webs sampled from three types of ecosystems (marine, freshwater, and lentic) were analysed. This is also the same dataset used in Chapter 5. Four most well-studied centrality metrics were applied to these food webs, including degree, betweenness, closeness and eigenvector centrality. For each species, the node rankings based on the four centrality metrics were compared in order to see their similarity. Specifically, this research aims to assess the effectiveness of different centrality metrics in differentiating species with different ecological and functional properties. Furthermore, food web robustness was measured by simulating species removal based on the four centrality metrics. Previous chapters found a strong correlation between robustness (targeted high degree species removal) and the relative core size. Here, the analysis was expanded to study the correlation among additional food web properties (i.e., proportion of basal / intermediate / top-level species, maximum trophic height, connectance, relative core size, etc.) with food web robustness under the four centrality targeted removal. The research presented

in this chapter provides a better understanding on centrality measurements to ecological networks, and the results can help identify key species from multiple perspectives.

6.2 Methods

6.2.1 Cumulative node centrality

Consider a graph, whereby each node has a certain centrality value c_j, and the nodes are ranked in descending order of centrality value. The cumulative centrality was proposed as the sum of all previous ranked centrality values and normalised by the sum of centrality of all the nodes:

$$C_n = \frac{\sum_{n'=1}^{n} C'_n}{\sum_j C_j} \tag{6.1}$$

where n is the centrality rank of the node.

The cumulative indice is useful when the purpose is to compare the data below or above a certain numerical threshold, for example, the proportion of nodes with degree over a certain value [199, 200]. Here the cumulative centrality was used to illustrate the distribution of centrality indices across all the nodes. For example, if the cumulative centrality reaches its maximum, it indicates that afterwards the nodes have centrality indices of zero. On the other hand, if the curve has a constant slope within certain segment, it means the nodes within the range have identical centrality indices, which also indicates the ineffectiveness of certain centrality metrics. Therefore, the cumulative centrality gives a clear view on the effectiveness of the particular centrality ranking. In order to compare whether different centrality metrics would generate the similar node ranking sequences, the Spearman correlation coefficient was applied in a pairwise manner.

6.2.2 Comparing robustness under different node removal sequences

Robustness was measured under the decreasing order of four centrality ranking sequences (degree, betweenness, closeness and eigenvector). Results were compared against random removal (1000 iterations for each food web). Existing research [17, 82, 138] and previous chapters in this thesis showed that targeted degree removal can break down food webs much faster than random removal. Therefore, if the robustness of any other types of centrality based removal is also significantly smaller than that of random removal, it can reflect the specific centrality ranking is also able to identify the species that are important to maintain the food web structure. Otherwise, if certain targeted removal generates similar results with random removal, it indicates that the related centrality metric is not effective in identifying key species that are responsible for maintaining the network structure.

6.3 Network properties and centralities for food webs

The 53 food webs were summarised with their network properties shown in Table A2. The size of the food webs ranges from 12 nodes to 92 nodes, with number of links ranges from 39 to 1644. The proportion of basal / intermediate / top-level species also varies hugely, and Table 6-A shows the network properties of the 8 selected food webs. Ythan Estuary food web only contains 4.3% of the basal level species, while for Broadstone Stream food web, $\%B$ is 66.7%, and for Coneyglen Burn food web, $\%B$ is 77.3%.

Table 6-A: A summary of the 8 food webs, including site name, which ecosystems they belong to, web size (N), link numbers (E), proportion of basal (B) / intermediate (I) / top-level species (T), connectance (C), and maximum trophic height (max TH). Food webs were ordered based on ecosystem types (1: freshwater streams; 2: lentic; 3: marine).

site	N	E	B	I	T	C	max TH
c4 [1]	65	422	0.492	0.154	0.354	0.100	3.599
d4 [1]	52	263	0.596	0.019	0.385	0.097	2.815
Broadstone Stream [1]	27	111	0.667	0.296	0.037	0.152	3.700
Skipwith Pond [2]	36	338	0.056	0.889	0.056	0.261	6.779
Tuesday Lake 1984 [2]	50	262	0.500	0.480	0.020	0.105	5.315
Tuesday Lake 1986 [2]	51	230	0.569	0.392	0.039	0.088	4.510
Ythan Estuary [3]	92	414	0.043	0.630	0.326	0.049	6.517
Benguela [3]	29	186	0.069	0.862	0.069	0.221	6.869

6.3.1 Distribution of node centrality metrics

Four most widely used centrality metrics were applied to the 53 food webs (degree, betweenness, closeness and eigenvector centrality, Figure 6.1). Nodes were ranked by the descending order of centrality metrics respectively, and normalised by the sum of centrality metrics for all the nodes.

In Figure 6.1, betweenness shares the most disparity with the rest of the three centrality metrics, as the curve is highly skewed towards the first few high betweenness nodes. Most of the food webs contain a large proportion of nodes with zero or approximately zero betweenness, especially for Ythan Estuary, d4, and c4. Networks always contain certain proportion of zero betweenness nodes which are not located on any shortest path. For example, in a star network, where all the nodes are connected to a central node, only the central node has betweenness larger than zero whilst all the other nodes have zero betweenness. In a chain network, the two nodes at the end of the chain always have zero betweenness. The proportion of zero betweenness nodes could increases significantly if the network becomes directed, because nodes at the start or the end of each directed chain will always have zero betweenness [201], for example, the resource species and the

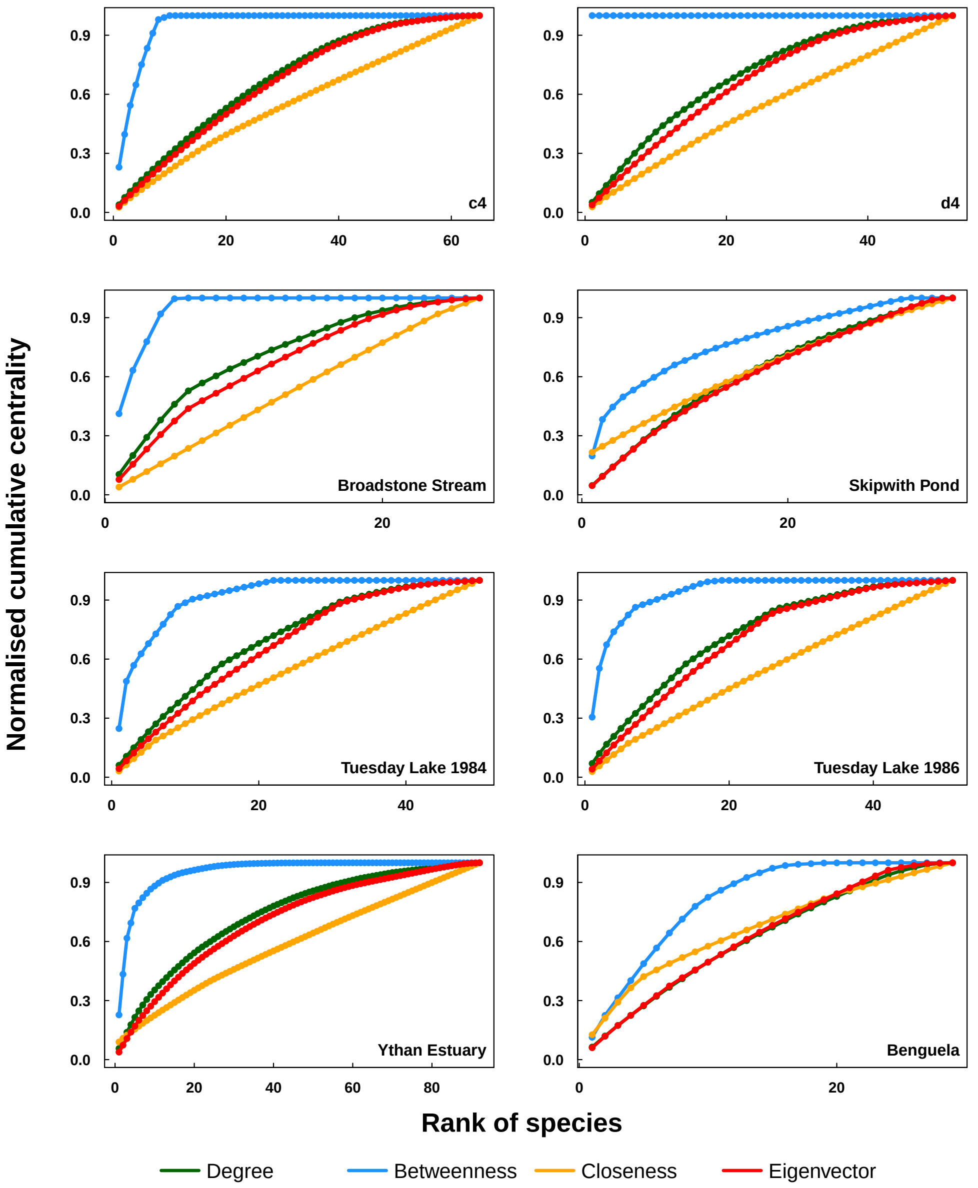


Figure 6.1: Normalised centrality ranking, including degree, betweenness, closeness and eigenvector centrality of 8 food webs. Nodes are ranked by the decreasing order of degree ranking, whose indices are shown on x-axis. Y-axis is the normalised accumulative centrality of each web. Food web names are shown at the bottom right corner.

top level predators that located at both ends of the food chain.

On the other hand, closeness centrality shows the most straight cumulative curve for most of the food webs, indicating that many species have identical closeness centrality values (except for Benguela food web, within which the top four nodes have much higher closeness). Therefore, both betweenness and closeness are not effective metrics to differentiate nodes. Degree centrality is effective in differentiating species as the cumulative curve increases sharply for the highest ranking species and grow steadily for nodes with lower number of links. Eigenvector centrality behaves similarly with degree centrality.

6.3.2 Implications of different centrality metrics to food webs

For each node, the variations of node centrality metrics were examined by comparing across the centrality indices C_j. The centrality metrics was normalised by the peak value (Figure 6.2 to 6.4). Each sub-figure (e.g., Figure 6.2**a**) consists of four panels (column **I** to **IV**), which illustrates the aforementioned different centrality indices. A clear distinguishability in the colour gradient indicates each node has different centrality value from each other. Based on this criteria, both betweenness (column **II**) and closeness (column **III**) centrality centralities are ineffective measures, because most nodes have either similar or identical centrality values.

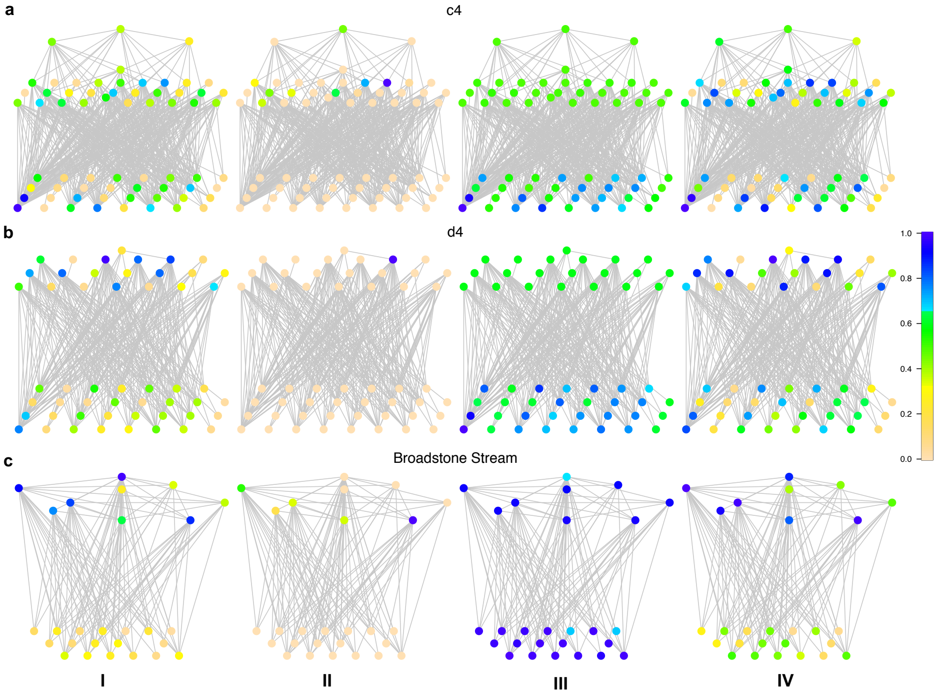


Figure 6.2: Centrality measurements were anlayzed on three stream food webs: **a** - Millstream control web (c4), **b** - Millstream drought web (d4), **c** - Broadstone stream. Species were coloured based on different centrality rankings, including **I** - Degree centrality, **II** - Betweenness centrality, **III** - Closeness centrality, **IV** - Eigenvector centrality.

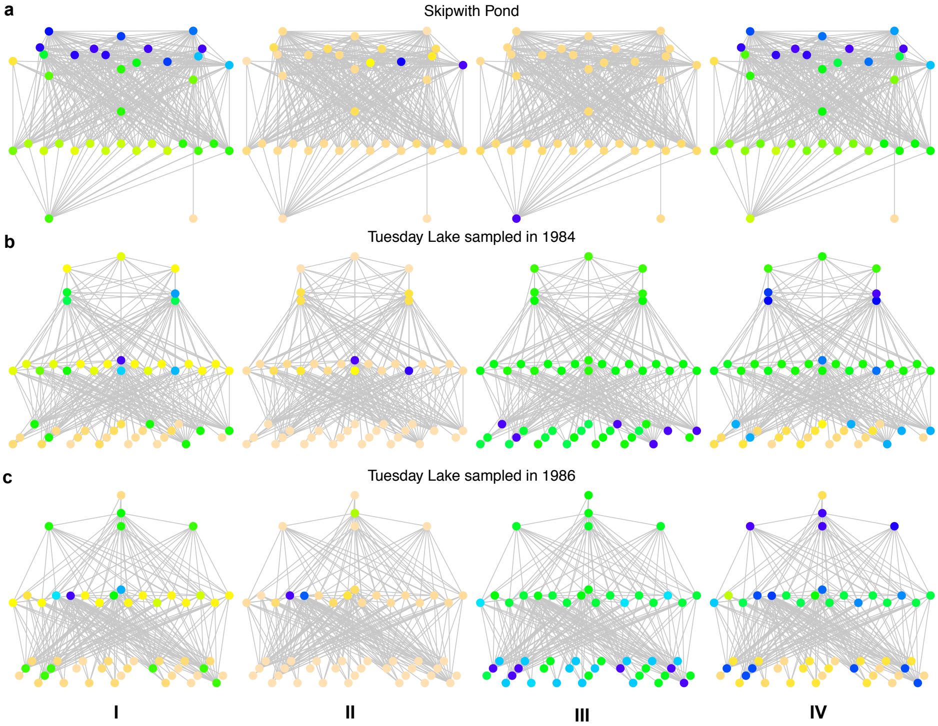


Figure 6.3: Centrality measurements were anlayzed on three lentic (e.g., lakes or ponds) food webs: **a** - Skipwith Pond, **b** - Tuesday Lake sampled in 1984, **c** - Tuesday Lake sampled in 1986. Species were coloured based on different centrality rankings, including **I** - Degree centrality, **II** - Betweenness centrality, **III** - Closeness centrality, **IV** - Eigenvector centrality.

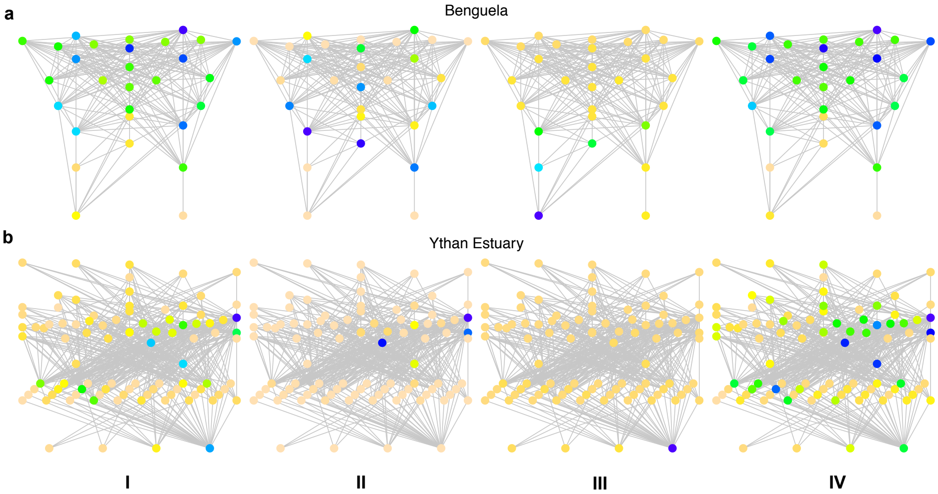


Figure 6.4: Centrality measurements were anlayzed on two marine food webs: **a** - Benguela, **b** - Ythan Estuary. Species were coloured based on different centrality rankings, including **I** - Degree centrality, **II** - Betweenness centrality, **III** - Closeness centrality, **IV** - Eigenvector centrality.

On the other hand, nodes were evenly distributed based on degree and eigenvector centralities (Figure 6.2 to 6.4). The ranking sequence of degree and eigenvector centrality were also similar with each other (based on the colour gradient). The results echo the fact that eigenvector centrality is a measurement of global degree centrality, since linking with a relatively higher degree nodes is able to add positive weight to the indices of its eigenvector centrality. Closeness measures how close a node to every other node, and therefore nodes in food webs with lower trophic height are expected to have a higher closeness. Results demonstrated that relatively low closeness centrality was observed in Skipwith Pond, Ythan Estuary, and Benguela food webs. Those webs also have a relatively high maximum trophic level (shown both in Table 6-A and in the Figure 6.3**a**, Figure 6.4). The trophic level plot for all the 53 food webs with nodes coloured by centrality indices can be found in Appendix Figure A2.

The Spearman correlation (ρ) between two centrality metrics highlighted the highest similarity between degree centrality and eigenvector centrality among all the food webs (Figure 6.5). Betweenness centrality also correlates strongly with degree and eigenvector centrality in stream food webs (c4, d4, and Broadstone Stream). Among the lentic food webs (Skipwith Pond and Tuesday Lakes), the correlation between betweenness and degree/eigenvector centrality is also strong, but slightly less than that among stream food webs. Closeness centrality, within all the food webs, correlates very weakly or even negatively with the other three indices. The correlation analysis for all the 53 food webs can be found in Appendix from Figure A3.

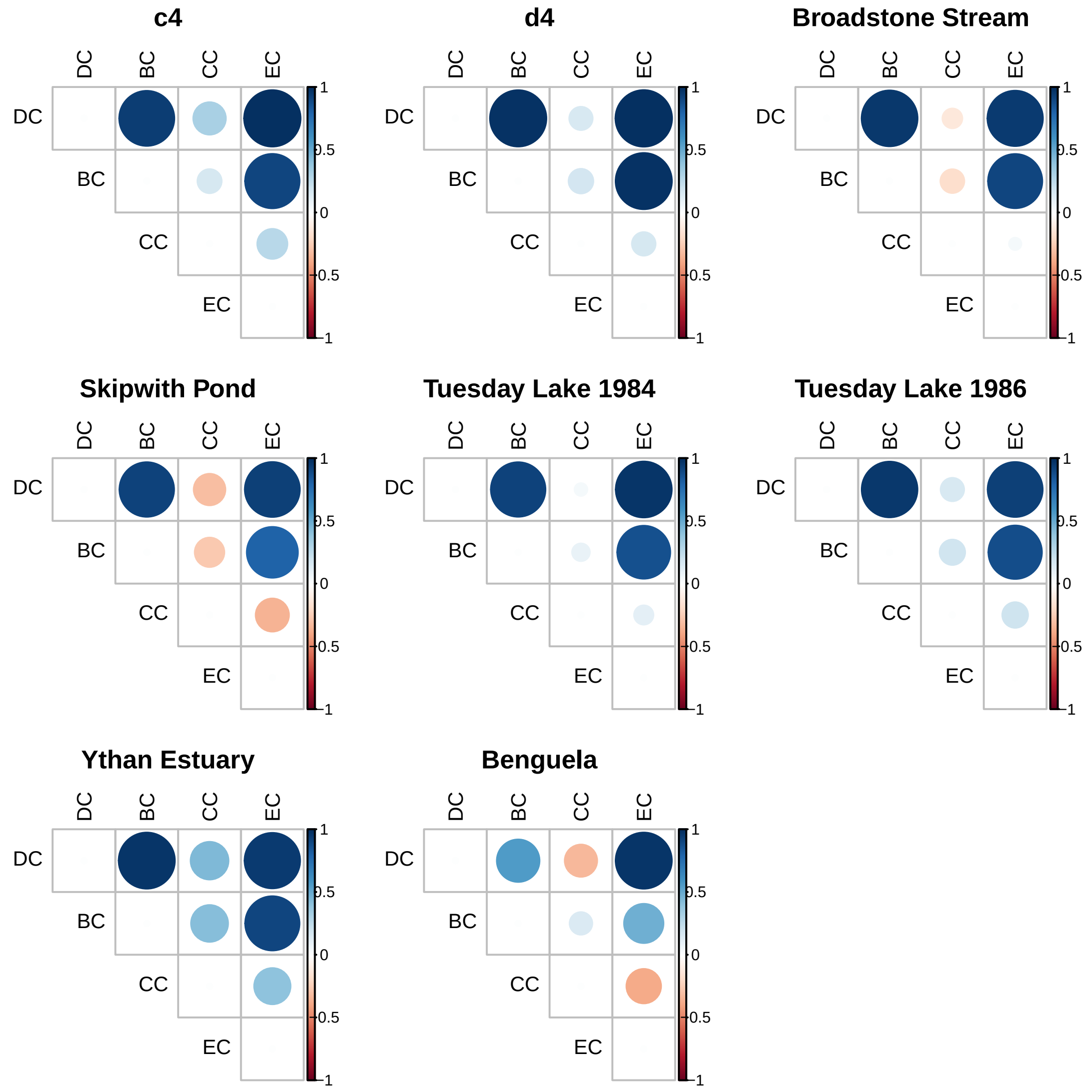


Figure 6.5: Comparisons of the correlation between centrality metrics using the Spearman correlation coefficient. Colour and size denotes the degree of covariation with larger blue ($\rho = 1$) as the most similar and small red ($\rho = 0$) as the least similar.

6.4 Robustness - removal sequences

In food webs, cascade failures happen when the extinction of certain species (simulated node removal in food webs) results in secondary extinction of other species. Here secondary extinctions only occur when species loose all its resources (Details in Section 2.4).

6.4.1 Comparison between targeted removal and random removal

All food webs showed a significant decrease in robustness under degree removal compared to that of random removal, with only Benguela food web showed a very slightly decrease in degree removal (Table 6-C). This finding is linked with Benguela food web has a much higher relative core size (0.793, Table A1), whilst other food webs have a relative core size that only ranges from 0.228 to 0.528. Betweenness is not an effective metric in detecting topologically important species, given targeted betweenness removal generates similar magnitude of secondary extinction compared to random removal. Even for some food webs, targeted betweenness removal generates less secondary extinction than random removal (Broadstone Stream, Skipwith Pond). Targeted closeness centrality removal is only effective in specific cases (Skipwith Pond, Ythan Estuary, and Benguela), with corresponding robustness much smaller than any other centrality removal as well as random removal. This can be explained by the fact that closeness centrality is able to identify basal level species accurately when the proportion of basal species is small (Figure 6.2 to 6.4, column **III**). Indeed, a small proportion of basal species (Skipwith Pond, $B\% = 0.056$, Benguela, $B\% = 0.069$. Table 6-A) increases the weight of dependence from higher level species on them, therefore the removal of any basal species can cause a severe collapse on the food web structure. The robustness under eigenvector removal is either the same or slightly higher than degree removal, as eigenvector centrality shows the strongest correlation with degree centrality (Figure 6.5). For robustness under the four centrality ranking removal of all the 53 food webs, please refer to Table A3.

6.4.2 Secondary extinction gradient through species removal

Figure 6.6 illustrates the common way of describing food web secondary extinction against simulated species removal based on centrality rankings. Each centrality is represented by its own respective curve which shows the cumulative secondary extinction against the proportion of species removed. As shown previously, degree and eigenvector

Table 6-C: Robustness under four centrality ranking sequences. Nodes were ordered in the decreasing order of centrality indices: Degree (DC), Betweenness (BC), Closeness (CC), and Eigenvector (EC) centrality. Results were compared with random removal, which run 100 iterations for each food web.

Site	DC	BC	CC	EC	Random (Avg.)
c4	0.323	0.369	0.323	0.338	0.471
d4	0.212	0.365	0.346	0.250	0.471
Broadstone Stream	0.185	0.500	0.407	0.222	0.463
Skipwith Pond	0.306	0.500	0.028	0.333	0.380
Tuesday Lake (1984)	0.240	0.160	0.280	0.200	0.460
Tuesday Lake (1986)	0.098	0.314	0.314	0.157	0.436
Ythan Estuary	0.049	0.239	0.011	0.054	0.376
Benguela	0.310	0.241	0.034	0.276	0.330

centrality are effective in achieving a more rapid rate of network collapse than the other two centrality measures. This is certainly true for d4, Broadstone stream, and Tuesday lake food webs. In the c4, Skipwith Pond, Ythan Estuary, and Benguela food webs, closeness centrality was as effective as degree and eigenvector centrality removal.

Yet, for Benguela, the random removal generates a similar level of secondary extinction compared to all the four types of targeted removal (Figure 6.6 and Table 6-C). Benguela food web is of a homogeneous nature which does not contain *bottleneck* species [19, 113] (i.e., the species that is responsible to connect different subgroups together), therefore it shows similar robustness under random removal and high centrality node removal. Food webs in which targeted removal resulted in a much smaller robustness are heterogeneous networks containing bottleneck species, as the removal of which can either break the network into small subnetworks or result in a larger number of isolated species [42]. In almost all of the cases, betweenness is not as effective as the other centrality measures and in the case of Broadstone stream and Skipwith Pond, it performs worse than the average of the random removal process (Figure 6.6 and Table 6-C).

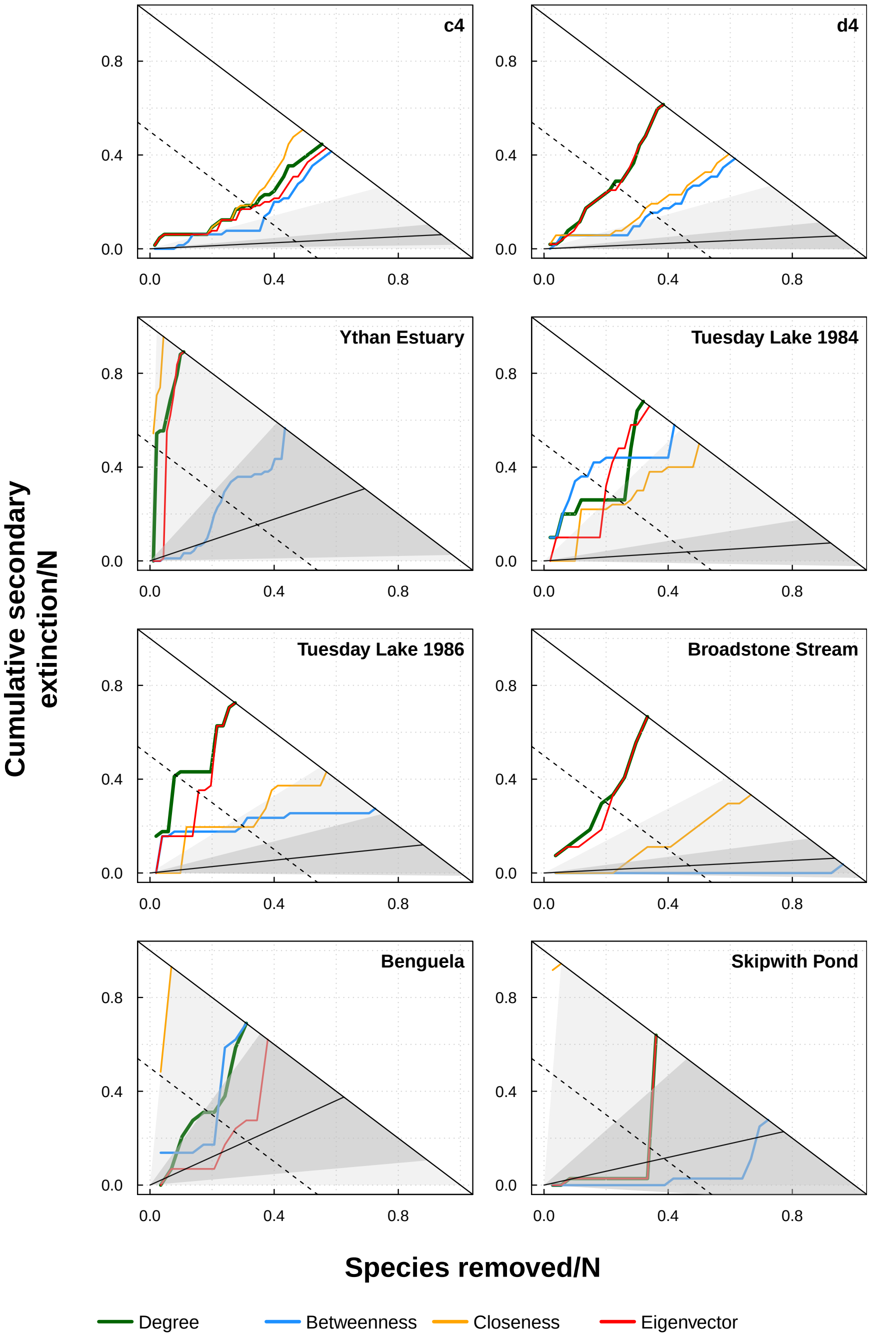


Figure 6.6: Robustness curve for eight food webs. Secondary extinction is shown on y-axis. Different removal strategies are shown in different colours. The diagonal line represents complete extinction of the network, whilst the parallel dashed line represents 50% species loss. The average random robustness is indicated by the black line. The grey shade envelopes the range of obtained from random removal, with the darker grey shade representing the one standard deviation from the average random robustness.

6.5 Correlation between centralities and network properties

Robustness was used under different centrality measurements to analyse how important species would influence the overall network structure, which is followed by the Pearson correlation in a pairwise manner between the different robustness and corresponding network properties (Figure 6.7).

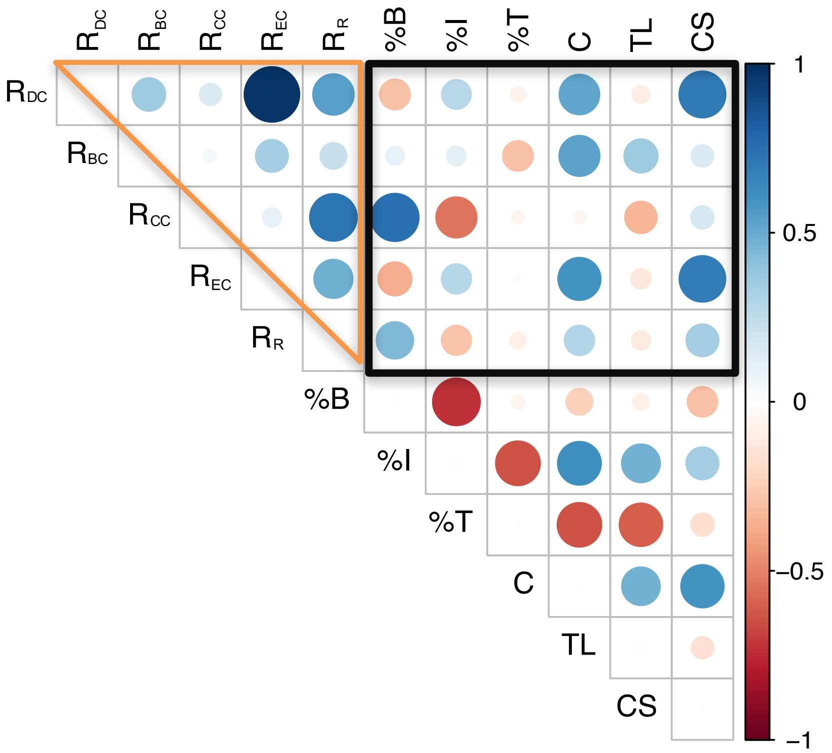


Figure 6.7: Correlation analysis between different robustness measurements and food web properties on the 53 food webs. Different centrality metrics are used for simulating targeted removal when measuring robustness (R). $R_{\mathbf{DC}}$ is based on degree ranking, $R_{\mathbf{BC}}$ is based on betweenness ranking, $R_{\mathbf{CC}}$ is based on closeness ranking, $R_{\mathbf{EC}}$ is based on eigenvector ranking, $R_{\mathbf{R}}$ is based on random ranking. **C** represents connectance, **TL** represents the maximum trophic level, and **CS** represents the relative core size. The resulting correlation ranges from -1 to 1, where 1 (dark blue) indicates a perfect positive correlation, and -1 (dark red) indicates a perfect reverse correlation.

Those highlighted in the black box in Figure 6.7 represent the correlation between robustness and network properties. Specifically, the proportion of basal level (B%) species was found to be strongly correlated with robustness under closeness based target removal, labeled as R_{CC} (dark blue, $r = 0.744$). This indicates that the smaller proportion the of basal species present in a food web, the more certain that the basal species are ranked top in the closeness centrality indices, which consequently leads to a smaller R_{CC}. In other words, $R_{\mathrm{CC}} \approx \mathrm{B}\%$, as $R_{\mathrm{CC}} = N_{\mathrm{rm}}/N$ and $\mathrm{B}\% = N_{\mathrm{basal}}/N$), and in this case, the removed nodes (N_{rm}) are almost the same with basal species (N_{basal}). The relative core size (CS) and connectance (C) both have strong correlations with degree ($r = 0.704$, and $r = 0.526$ respectively), echoing the findings in Chapter 5.

Those highlighted in the orange triangle represent the correlation among different robustness metrics (Figure 6.7). Strongest correlation was found when robustness under degree removal was compared with robustness under eigenvector removal ($r = 0.978$). Robustness under betweenness removal (R_{BC}) failed to show a relatively strong correlation with any other robustness metrics.

In addition, the correlation analysis showed a number of network properties of food webs also have strong correlations. For example, the maximum trophic level (TL) was found to negatively correlate with the proportion of top-level species, meaning that the higher level a food web reaches, the fewer top-level species it could have.

6.6 Summary

Whether a food web shows homogeneous or heterogeneous structural pattern can strongly affect its robustness under species removal. For example, homogeneous networks lack strategically located nodes known as bottlenecks that link densely connected subgroups together. Therefore, homogeneous networks show similar robustness under both random and targeted removal (e.g., Benguela food web) [113]. Whilst in heterogeneous networks, centrality metrics help to identify important species which are the key to maintain robustness, therefore, the removal of these high centrality species can cause much more impact than random removal.

Degree centrality identifies key species with the most interactions, therefore the remove of which can cause the most devastating effect on the network, resulting in the smallest robustness. Betweenness centrality is less effective given that food webs tend to contain many zero betweenness nodes. This is caused by a large proportion of basal and top-level species which never lie on the shortest paths between nodes. Food webs in general have large closeness centrality indices for all the nodes, given its densely connected features and small average shortest path lengths [60, 202]. If a food web contains very few basal species, and it has links to the majority of the higher level predators, then those basal species are considered to be located in the "center" of the network with the highest closeness. Therefore, closeness is extremely effective to identify key basal species. This leads to the strongest correlation between the proportion of basal level species and robustness based on targeted high closeness node removal.

The results presented in this chapter filled the void in current research, as not only degree centrality was used to identify important species in food webs [17, 80, 192, 203], other centrality metrics were also used to explore key species from novel perspectives. These results demonstrate that whilst the aspects of identifying key species are multi-dimensional and complex [20, 196, 204], the centrality metrics can shed light on better understanding the unique topological feature of food webs.

Chapter 7

Conclusions and Future work

7.1 Conclusions

The key contribution of this research is the discovery of the core substructure presented in food webs, which is responsible in governing food web robustness under various kinds of perturbations such as environmental stressors [36, 81, 103] or management practices [84]. Network properties related to the core substructures can help unveil hidden adaptive behaviours of food webs under perturbations, which explain the species level properties (i.e., species extinction or invasion, abundance or biomass changes, connectance changes, etc.) and their connections with food web functional properties (i.e., robustness).

In the first part of the thesis (Chapter 3), the results show that the freshwater food webs under drought stress conserve their richly connected core structure and the relative core size despite sharp decreases in their biodiversity and biomass production [36, 50]. The underlying mechanism is a movement of core species to the periphery, which preserves the relative core size. As a consequence, food web robustness is also maintained. This new way of analysing food webs provides a more comprehensive understanding on the impact of perturbations at the substructure level, which leads to new directions for future research on community response under climate change.

In the second part of the thesis (Chapter 4), the results show the importance of food web substructure in relation to how different farming practices would affect ecosystem biodiversity and robustness. New farming practices on Genetically Modified Herbicide Tolerant (GMHT) crops are claimed to have long-term influence on sustainable agriculture as well as the wider ecosystem stability [177, 205–207]. By studying 502 farmland food webs across the UK on different crop types, the results demonstrate that despite GMHT management exerts significant effect on species level properties, both substructure level properties and the robustness are not significantly affected by it. On the contrary, crop switching causes significant disturbances, which alters both the species level, substructure properties, and robustness. Furthermore, the most common species appear more frequently in the core structure, and the conservation of a core structure across all the agricultural food webs is associated with the fact that the species composition within the core is much more stable than periphery. The empirical food web studies on both freshwater and agricultural ecosystem (Chapter 3 and Chapter 4) imply the stabilising effect of a core structure, as both studies show that species composition is more stable in the core, which governs the most of the energy flow and interactions within the system.

The underlying reasons why a core structure can act as an indicator of food web robustness is further elaborated in Chapter 5. The presence of a core acts as a stabilising factor for networks: a large core is often associated with more alternative pathways among nodes [54, 55, 170, 171]. As a result, the core can mitigate the effects of external perturbations. By incorporating over 50 food webs from different ecosystems, the research reported in this thesis is one of the first analysis of the importance of the core structure in food webs. Specifically, the relative size of the core among food webs strongly correlates with their robustness. Connectance was previously shown to be positively correlated with robustness [17, 43, 79]. Here the relative core size is shown as a stronger indicator of robustness compared to connectance. This is true especially for small food webs, as the correlation between relative core size and robustness was consistently strong whilst the correlation between connectance and robustness dropped significantly when smaller webs were taken into consideration. Indeed, most of the redundancy was conserved within the core, and the core redundancy covaries with the relative core size. Results demonstrated that the distribution of redundancy is important in characterising food web robustness, while connectance as a network level measurement, cannot capture this substructure level properties.

Further analysis of the food webs is performed in Chapter 6, and it is done so from the perspective of robustness in the face of targeted removals based on different centrality rankings and random removals. The random extinction sequence removal indicates the average resilience level of a food web, and the decreasing order of degree sequence removal measures the worst-case scenario of a food web under perturbations [17, 82, 84]. Existing works have evaluated how various types of targeted node removal would result in different robustness [71, 153, 191], yet the underlying reasons why certain centrality ranking works better than others was not fully explained. Here different removal sequences based on various centrality indices are examined on over 50 food webs and more importantly, the implication of each centrality ranking was linked with specific species properties. For example, species with high closeness centrality are mostly basal species from the bottom level of the food webs, indicating that as the resource provider of the food web, basal species locate in the central of the network thus can provide energy to the higher level species efficiently. However, not all centrality metrics can effectively profile and distinguish species, for example, betweenness only highlights a small proportion of species to have high centrality, whilst for the rest of the nodes, betweenness cannot clearly distinguish them based on centrality indices.

This research also examined how the ranking of each species varies according to different centrality metrics. High degree species is shown not only to be able to break down the food web in the fastest way [191], but also able to distinguish nodes more clearly based on the spectrum of the degree distribution. Node ranking based on eigenvector centrality shares the highest similarity with degree ranking, whilst closeness centrality shares almost no similarity with all the other centrality rankings. This finding contradicts previous work showing that closeness centrality strongly correlates with degree centrality on random graphs and non-ecological networks [153], indicating food webs in general possess unique topological structures. However, closeness centrality is strongly associated with basal species, and as such, the targeted removal of high ranking closeness species is able to break down the food web. In summary, it is vital to select the most appropriate centrality indices when analysing food webs.

The network analysis methods presented in this thesis can be applied to other networks from three aspects. First, if the focus is to evaluate the external perturbations on ecosystems, it is possible to compare the network substructure properties before and after the perturbation in order to quantify the level of impact using the methodology presented in the thesis. Second, if the focus is to study the food web structural robustness, different node removal methods can be applied and its effectiveness can be evaluated. Finally, the network analysis methods are generic and can be applied to other non-ecological networks, given that the nodes representing the components in a system and the links representing their interactions. In summary, the network analysis methods provided here explored empirical networks from a new substructure perspective, and the focus on exploring the self-organising behaviour of the node-link interactions provides insights on better understanding on system robustness.

7.2 Future work

Food webs have been shown to possess a richly-connected core structure, regardless of their type and residing ecosystems (see Chapters 3 to 6). In this thesis, food webs were studied in their unweighted form and it is important to further expand current research to consider weighted food webs. Yet it is not immediately obvious what approach of weighting makes sense. In general, if the weight of a link is w, then an exponent

w^{α} ($\alpha = \{0, 1\}$) has been shown to be an effective way to adjust the contribution of weights [208], so that a balance exists between distinguishing the number of links and the weight of links. An emphasis on the weighting approach ($\alpha = 1$) will allow one to better understand the key energy flow in a food web, but it will risk obscuring the importance of diversity (i.e., the number of links with significantly lower weights) and suppress subgraph structures. In terms of methodology, there are various ways to weight a food web, for example, nodes can be weighted by species body size (M), abundance (N), or biomass ($B = f(M, N)$) [209, 210], and links can be weighted by biomass or energy flow [211]. From a dynamic perspective, a weighted food web contains more detailed information than unweighted one, as the changing of weight might indicate how climate change would affect the interaction strength between species. For example, the production of dominant invertebrate species in a post-flood stream increased significantly, which directly resulted in increasing in the trout biomass, as trout mainly feeds on those invertebrates [211]. Since the perturbation mostly affects the link weight, the unweighted version of the food web shows almost no change on food web structure before and after flood, therefore cannot detect those dynamics in the interaction strength changes. Future research can focus on investigating how the core structure would change with or without node/link weight. The inclusion of weight focuses on the energy flow instead of the actual linkages, which might be able to uncover previously unfound community responses when perturbation happens. With advances in community detection methods being introduced to analysing biological networks [169, 212], the substructure analysis on food webs can be enlightened from new perspectives, in order to better understand the community response under environment stressors.

Real-world complex networks are highly dynamic not only because they grow [94, 213, 214], but also the links can re-wire in face of perturbations [215–219]. For example, biological networks, such as the protein-protein interaction networks are believed to possess high degree of adaptability under many iterations of the evolutionary process [220]. In food webs, despite various types of both biotic or abiotic perturbations, they are able to persist by adapting to the changing environment, i.e., establishing new predatory links in order to prevent them from extinction [1, 36, 50, 81, 168]. Simulated species removal based on adaptive behaviour has been applied to better evaluate food web robustness by re-wiring links according to theoretical modelling [14, 16, 77, 78]. However, the way adaptation was modelled is either body-size based, which assumes larger predators are all expected to switching to prey within a size range, or topological based, which assumes species from a higher level are expected to prey on the species from lower levels [16]. An important step to achieve realistic re-wiring modelling is to better understand how species would switching diet under real world observations. It is important to identify the more preferable links when adaptation occurs. A recent microcosms experiment showed functional redundant links between predatory invertebrates and its prey can help maintain the functions of heterotrophs under the stress of prey extinction caused by herbicide [221]. The predators mainly relied on consuming non-predatory invertebrates (occupying up to 80% of the proportion in their diet) under control treatment, and after 8 weeks of herbicide application, the sensitive invertebrates died out, forcing the predators to shift their diet towards consuming detritus (occupying 100% proportion in diet under the highest herbicide concentration). This finding suggests re-wiring to new resources only happens if the preferable type of resources died out.

Attaching new links can also be achieved based on preferential attachment from a topological perspective, for example, in social networks, if two nodes share many common neighbours, then the two nodes are very likely to be linked with each other in the future [222]. Similar ideas might be applied to food webs: a rewired link can be preferred than another if the selected new prey shares more similarity than the one that the predator lost.

A key feature of food webs is that they always contain certain level of redundancy in order to function properly, as redundancy not only maintains the ecosystem functionality, but also enhances certain functions that cannot be achieved without redundancy. For example, in a experiment where the abundance of a pest needs to be controlled by introducing its predator (ant species) into the ecosystem, the co-existence of different predators effectively reduced the abundance of the pest whilst only introducing one ant species did not perform effectively [223]. This can be explained by the multiple predator effects on their common prey. This study reveals that the synergistic effects within ecosystem act as an important factor when characterising the consequences from various stressors upon food webs, which can only be assessed through a network approach. Future research need to investigate the functional importance of redundant links, which can be applied to food web modelling and to better evaluate ecosystems.

Appendix A

Author's publications

Journal papers

1. [**Lu, 2014**]: Clare Gray, Donald J Baird, Simone Baumgartner, Ute Jacob, Gareth B. Jenkins, Eoin J O'Gorman, **Xueke Lu**, Athen Ma, Michael J. O. Pocock, Nele Schuwirth, Murray Thompson and Guy woodward "Ecological networks: the missing links in biomonitoring science" *Journal of Applied Ecology* 51 no. 5 (2014): 1444-1449.

2. [**Lu, 2016**]: **Xueke Lu**, Clare Gray, Lee E Brown, Mark E Ledger, Raúl J Mondragón, Guy Woodward and Athen Ma "Drought rewires the cores of food webs" *Nature Climate Change.* Published online on 09 May 2016 (doi:10.1038/nclimate3002).

Conferences attended

1. **The EcoNet2015 workshop in Umea Marine Science center, Sweden.** *June, 2015.* Presentation with title "Modeling food web robustness with adaptability behavior".

2. **Seminar at South Kensington, Imperial College, centre for complexity**

science. *November, 2014.* Invited talk with title "Self-organization and robustness in food webs"

3. **The Student Conference on Complexity Science (SCCS conference).** *August, 2014.* Presentation with title "Self-organisation and robustness in food webs".
4. **British Ecological Scientific Aquatic Group conference (BES-AG conference).** *July, 2014.* Poster with title "Exploiting link between food web robustness & network cohesiveness".
5. **Grand Challenge in Ecosystems and the Environment workshop (GCEE workshop).** *July, 2014.*
6. **Network group seminar.** *May, 2014.* Presentation with title "Exploiting self organisation and robustness in ecological networks".
7. **School of Electronic Engineering and Computer Science Post Graduate conference (EECS PG conference).** *July, 2013.* Poster with title "Exploiting link between food web robustness & network cohesiveness".
8. **Symposium workshop at Silwood Park.** *August, 2013* Presentation with title "Exploiting ecological networks to understand artificial properties".
9. **BES Annual General meeting at INTECOL Congress, Ecology: Into the next 100 years.** *August, 2013.*
10. **the Cambridge Networks Day (CNDay 2013).** *May, 2013.*
11. **Large Evolving Networks confernece**. *Feburary, 2013.*

Appendix B

Supplementary data and results

B.1 Z-score evaluation on the rich-club coefficient with random networks

In chapter 3, z-score is used to compare the rich-club coefficient of empirical food webs with random networks generated by the null model in order to show the significance of food web having certain kind of connectivity feature among high degree species.

Specifically, z-score (Eq. 2.6) is calculated as:

$$z_{\phi_i} = \frac{\phi_i - \frac{\sum_j^N \phi_{j_i}}{N}}{\sqrt{Var(\phi_{j_i})}} \tag{B.1}$$

where ϕ_{j_i} is the rich-club coefficient of the jth random graph at node i where the null model is run for N times. Figure A1 plots the comparison between ϕ_r and $\phi_{r_{null}}$ of each food webs.

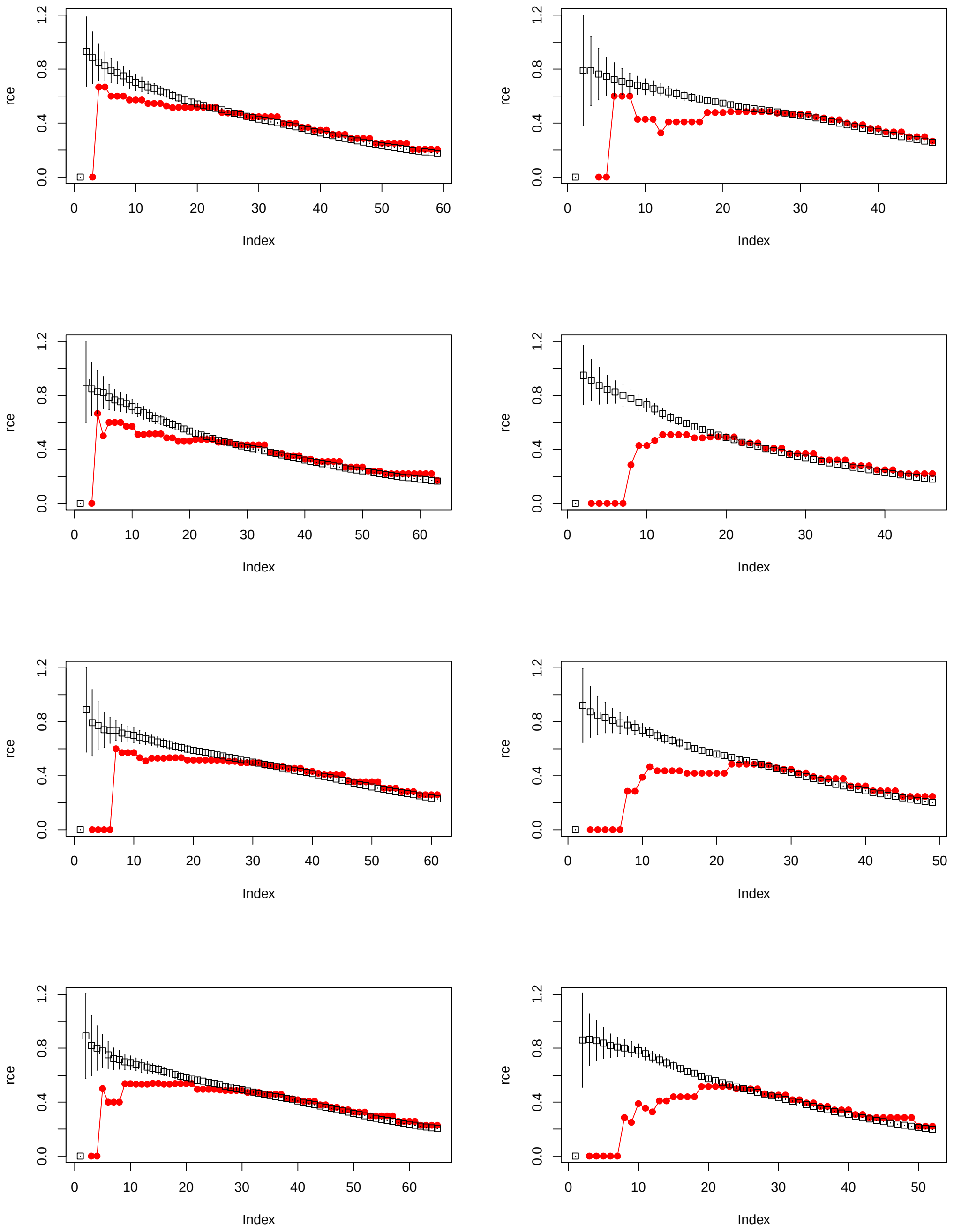


Figure A1: $\phi_{r_{\text{null}}}$ (black box plot, with mean and standard deviation) is compared with ϕ_r (red) for both control (left) and drought webs (right).

B.2 Core/periphery examples

Examples listed below show how the core/periphery concepts are applied to different networks.

World trade web

In the world trade web [224], which contains 160 involved countries (as nodes) and business relationship (as links), the core detected using the core profiling algorithm contains over 100 nodes. If link weight is included, the core only contains seven nodes. Unweighted network has an focus on the relationship among countries, therefore, the resultant core contains all the major EU countries and the US; while weighted network focuses on the business frequency and strength, resulting in the core with only world's top importers and exporters.

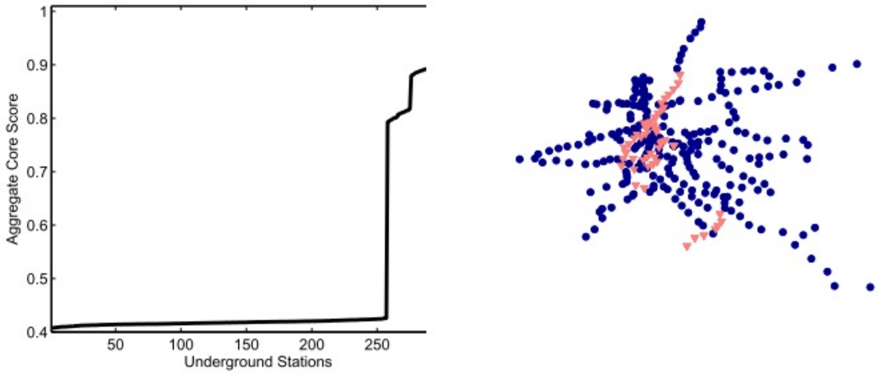
London underground network

In the London underground network, which contains 257 stations and lines between them (Figure A2**a**), there are sixty stations belong to the core, shown in Figure A2**b**. The core stations locate in central London and are connected to major train stations. For example, the King's Cross Station and the Waterloo Station rank top in the core nodes profiling, and they are also hubs for both trains and boats [57].

Brain network

Brain networks contain nodes as cortical and subcortical regions and links as functional relations among the nodes. Around 20 out of 120 brain regions are identified as core nodes. The core nodes are also identified as the least flexibility regions when learning tasks are changed one after another sequentially. Similar results are obtained under various type and intensity of training [58]. Results suggest the core is the most stable substructure under perturbations.

Table A1 summarises the significance of the core/periphery structure in above networks.



(a) (b)

Figure A2: (a) The increasing order of the importance of core nodes for all the tub stations. (b) Geographical plot of all the tube stations, with core highlighted in pink symbols [57].

Network type	size of the core N_c/N_{all}	Significance of core nodes
World trade web (weighted)	100/160	US and most European countries
World trade web (unweighted)	7/160	World's top importers and exporters
London tube	60/257	Located in central London, connected with major train/ferry stations
Brain network	20/120	Least flexibility compared with periphery nodes

Table A1: The significance of the core nodes in real-world networks.

Appendix C

Food web data

C.1 53 standard food web data

Site	Ecosystem	N	Connectance	Core size	R_{core}	robustness
Cuyahoga River	stream	12	0.319	0.833	0.955	0.333
Mahoning River	stream	13	0.367	0.692	0.797	0.385
Sandusky River	stream	13	0.231	0.615	0.861	0.231
Buck Creek	stream	18	0.191	0.444	0.726	0.167
Duddon main channel	pH stream	19	0.186	0.842	0.949	0.21
Duddon Beck	pH stream	20	0.27	0.8	0.888	0.4
West Branch Mahoning River 2	stream	20	0.335	0.7	0.831	0.4
Dargall Lane	pH stream	21	0.213	0.762	0.881	0.286
Mosendale Beck	pH stream	21	0.234	0.81	0.893	0.333
Allt na Coire nan Con	pH stream	22	0.184	0.727	0.89	0.318

Site	Ecosystem	N	Connectance	Core size	R_{core}	robustness
Coneyglen Burn	pH stream	22	0.114	0.227	0.55	0.091
Duck Creek	stream	22	0.275	0.591	0.796	0.273
Old Lodge	pH stream	23	0.251	0.87	0.926	0.348
Afon Gwy	pH stream	24	0.227	0.875	0.936	0.375
Lone Oak	pH stream	24	0.271	0.917	0.953	0.417
Paint Creek 2	stream	24	0.177	0.458	0.755	0.125
West Branch Mahoning River	stream	24	0.292	0.667	0.854	0.25
Hocking River 2	stream	24	0.247	0.583	0.818	0.208
Blanchard River	stream	24	0.274	0.583	0.791	0.292
Afon Hafren	pH stream	25	0.21	0.84	0.909	0.4
Broadstone	pH stream	25	0.275	0.72	0.809	0.44
Broadstone Stream	stream	27	0.172	0.444	0.74	0.185
Benguela	stream	29	0.233	0.793	0.935	0.31
Wrynose Pass Beck	pH stream	29	0.22	0.828	0.929	0.379
Hocking River	stream	29	0.25	0.621	0.851	0.241
Beagh's Burn	pH stream	30	0.204	0.867	0.936	0.267
Scioto Brush Creek	stream	30	0.262	0.567	0.78	0.233
Duddon Pike Beck	pH stream	35	0.229	0.771	0.851	0.371
Big Darby Creek	stream	35	0.231	0.486	0.74	0.257
Skipwith Pond	lake	36	0.286	0.528	0.725	0.306
Paint Creek	stream	38	0.186	0.395	0.693	0.184
Little Miami River	stream	39	0.254	0.513	0.747	0.256
Walhonding River	stream	39	0.262	0.513	0.734	0.256
Little Miami River 3	stream	39	0.238	0.513	0.76	0.231
Allt a'Mharcaidh	pH stream	40	0.204	0.65	0.772	0.4
Little Miami River 2	stream	40	0.236	0.45	0.827	0.275
Etherow	pH stream	44	0.218	0.523	0.613	0.386

Site	Ecosystem	N	Connectance	Core size	R_{core}	robustness
Hardknott Gill	pH stream	44	0.197	0.682	0.803	0.386
d2	Millstream	46	0.088	0.435	0.729	0.152
d1	Millstream	47	0.125	0.574	0.795	0.319
d3	Millstream	49	0.099	0.469	0.761	0.204
Tuesday Lake 1984	lake	50	0.106	0.4	0.682	0.24
Tuesday Lake 1986	lake	51	0.091	0.314	0.63	0.098
d4	Millstream	52	0.097	0.423	0.714	0.212
c1	Millstream	59	0.086	0.508	0.809	0.271
c3	Millstream	61	0.112	0.541	0.789	0.312
Narrator Brook	pH stream	61	0.201	0.508	0.656	0.328
c2	Millstream	63	0.081	0.492	0.784	0.254
c4	Millstream	65	0.1	0.492	0.762	0.323
Bere Stream	pH stream	65	0.222	0.692	0.778	0.431
Kennet	Kennet	83	0.131	0.325	0.552	0.265
Mill Stream	pH stream	87	0.218	0.598	0.705	0.425
Ythan Estuary	marine	92	0.049	0.228	0.57	0.044

Table A1: A summary of all 53 food webs, including food web name, which ecosystems (or groups) they belong to, web size, robustness, connectance, the relative core size and overall network redundancy. Food webs are ordered by the decreasing order of the network size.

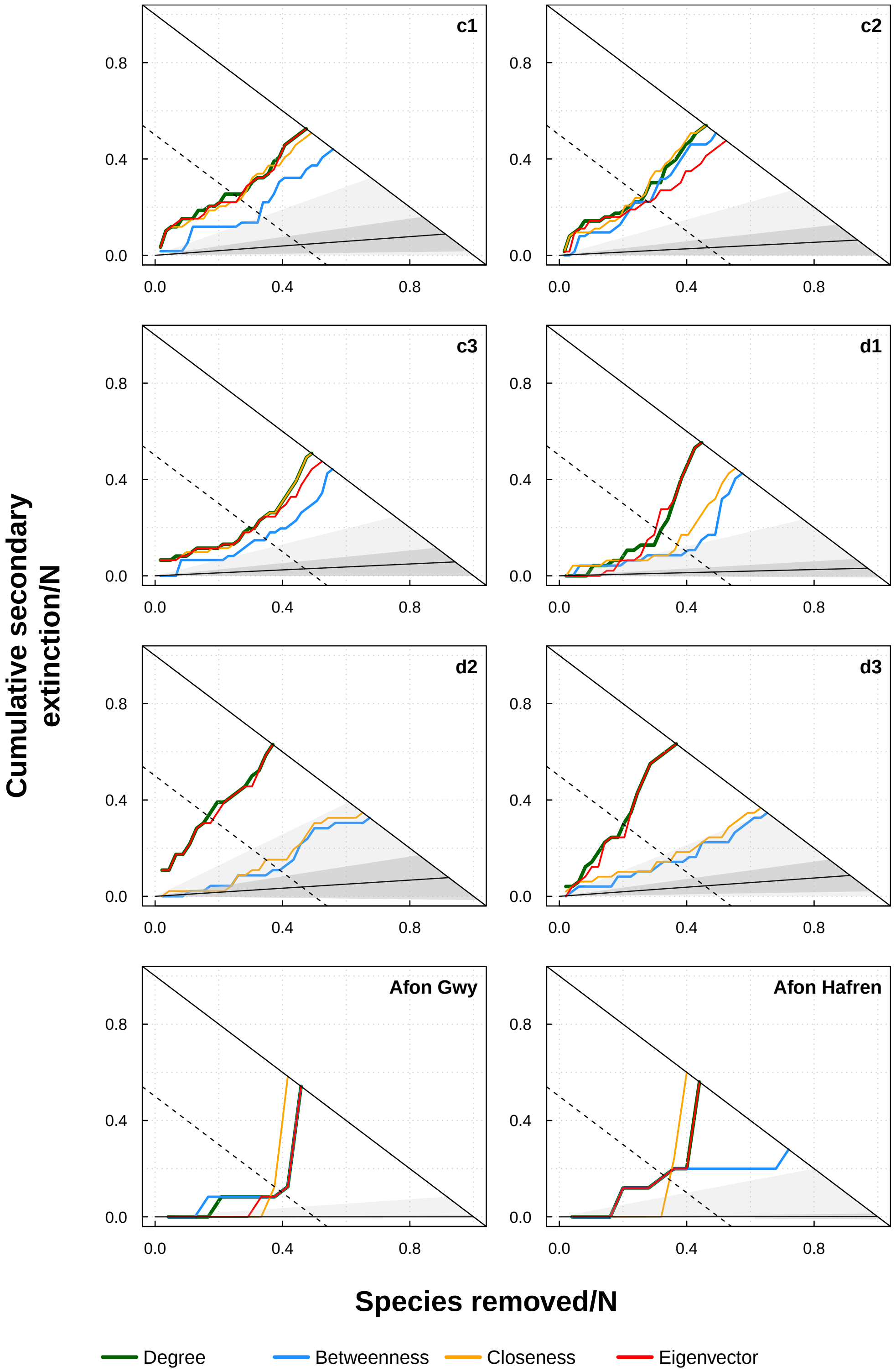
index	site	N	E	B	I	T	C	max TH
1	c1	59	300	0.492	0.102	0.407	0.086	3.308
2	c2	63	323	0.460	0.048	0.492	0.081	2.857
3	c3	61	418	0.492	0.098	0.410	0.112	3.254
4	c4	65	422	0.492	0.154	0.354	0.100	3.599
5	d1	47	277	0.553	0.021	0.426	0.125	2.636

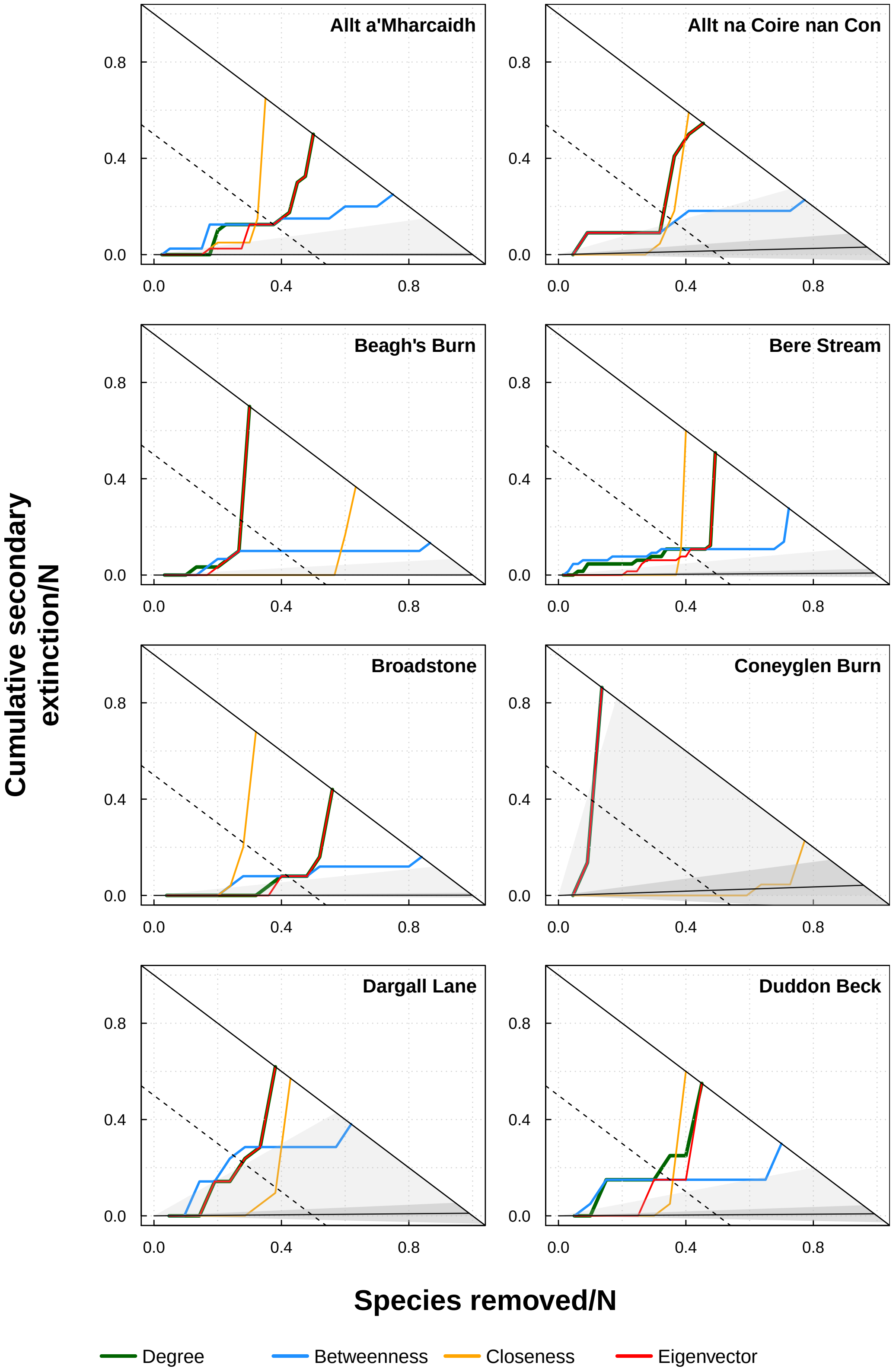
6	d2	46	186	0.652	0.022	0.326	0.088	2.769
7	d3	49	238	0.633	0.020	0.347	0.099	2.885
8	d4	52	263	0.596	0.019	0.385	0.097	2.815
9	Ythan Estuary	92	414	0.043	0.630	0.326	0.049	6.517
10	Tuesday Lake 1984	50	262	0.500	0.480	0.020	0.105	5.315
11	Tuesday Lake 1986	51	230	0.569	0.392	0.039	0.088	4.510
12	Broadstone Stream	27	111	0.667	0.296	0.037	0.152	3.700
13	Benguela	29	186	0.069	0.862	0.069	0.221	6.869
14	Skipwith Pond	36	338	0.056	0.889	0.056	0.261	6.779
15	Afon Gwy	24	131	0.417	0.500	0.083	0.227	3.626
16	Afon Hafren	25	129	0.400	0.400	0.200	0.206	3.353
17	Allt a'Mharcaidh	40	325	0.350	0.500	0.150	0.203	4.043
18	Allt na Coire nan Con	22	89	0.409	0.409	0.182	0.184	3.761
19	Beagh's Burn	30	184	0.633	0.300	0.067	0.204	3.845
20	Bere Stream	65	931	0.400	0.446	0.154	0.220	5.220
21	Broadstone	25	158	0.320	0.600	0.080	0.253	4.189
22	Coneyglen Burn	22	55	0.773	0.182	0.045	0.114	3.500
23	Dargall Lane	21	92	0.429	0.476	0.095	0.209	4.009
24	Duddon Beck	20	96	0.400	0.450	0.150	0.240	3.574
25	Duddon main channel	19	65	0.526	0.316	0.158	0.180	3.857
26	Duddon Pike Beck	35	274	0.371	0.457	0.171	0.224	4.259
27	Etherow	44	421	0.409	0.341	0.250	0.217	3.506
28	Hardknott Gill	44	375	0.341	0.614	0.045	0.194	4.711
29	Lone Oak	24	152	0.417	0.458	0.125	0.264	3.410
30	Mill Stream	87	1644	0.379	0.506	0.115	0.217	7.595
31	Mosendale Beck	21	99	0.476	0.429	0.095	0.224	3.965
32	Narrator Brook	61	738	0.311	0.525	0.164	0.198	4.718

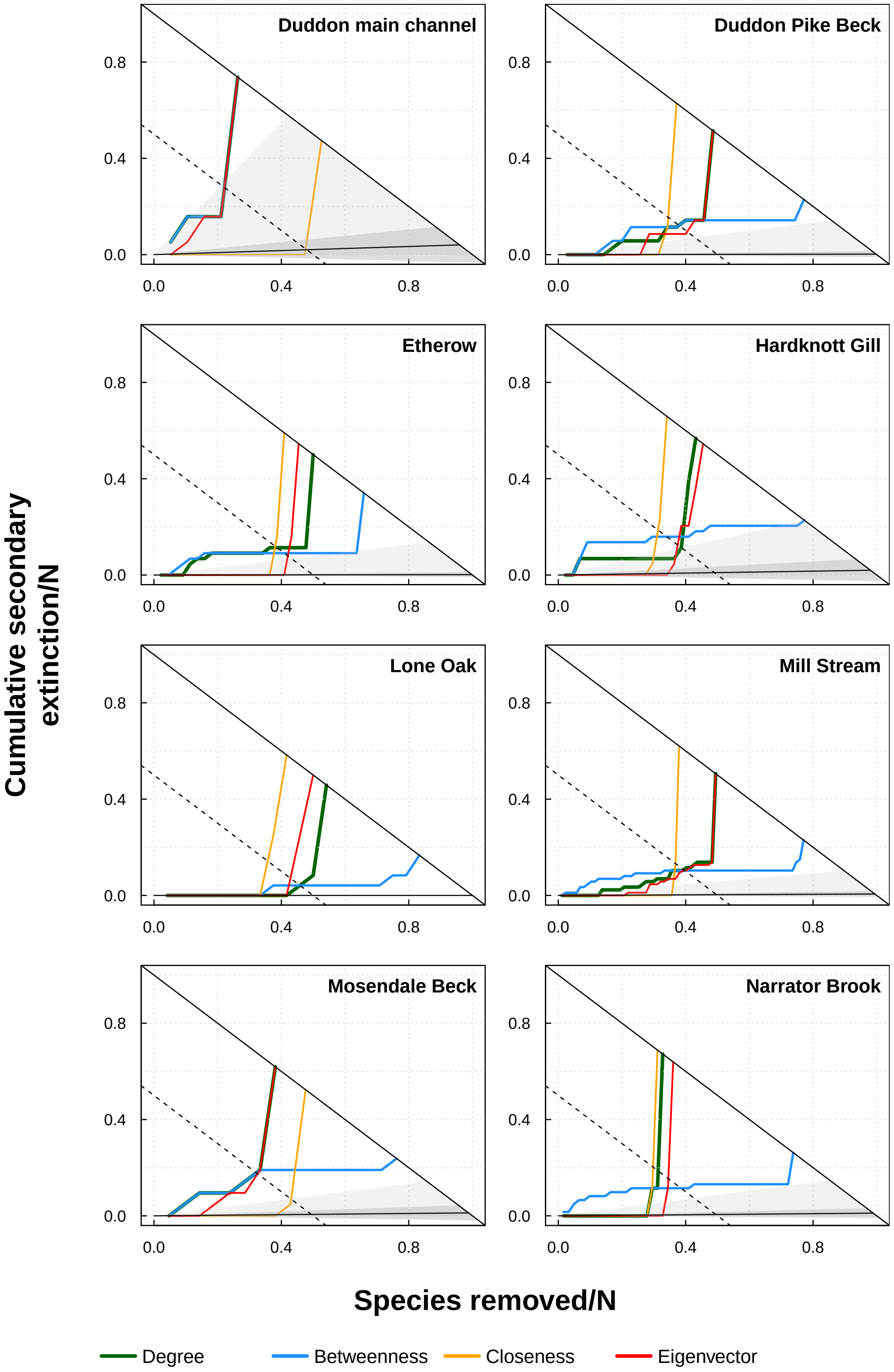
33	Old Lodge	23	129	0.522	0.391	0.087	0.244	3.926
34	Wrynose Pass Beck	29	183	0.448	0.414	0.138	0.218	3.802
35	Kennet	83	901	0.602	0.313	0.084	0.131	5.879
36	Scioto Brush Creek	30	236	0.533	0.433	0.033	0.262	7.885
37	Little Miami River	39	387	0.513	0.462	0.026	0.254	10.064
38	Little Miami River	40	378	0.600	0.375	0.025	0.236	9.020
39	Buck Creek	18	62	0.611	0.333	0.056	0.191	4.444
40	Big Darby Creek	35	283	0.571	0.400	0.029	0.231	8.461
41	Paint Creek	38	269	0.658	0.316	0.026	0.186	7.207
42	Paint Creek	24	102	0.667	0.292	0.042	0.177	5.171
43	West Branch Mahoning River	24	168	0.417	0.542	0.042	0.292	7.989
44	West Branch Mahoning River	20	134	0.450	0.500	0.050	0.335	6.766
45	Mahoning River	13	62	0.385	0.538	0.077	0.367	5.100
46	Walhonding River	39	399	0.564	0.410	0.026	0.262	9.306
47	Hocking River	29	210	0.517	0.448	0.034	0.250	8.349
48	Hocking River	24	142	0.500	0.458	0.042	0.247	7.266
49	Duck Creek	22	133	0.455	0.500	0.045	0.275	6.956
50	Cuyahoga River	12	46	0.417	0.500	0.083	0.319	4.949
51	Sandusky River	13	39	0.462	0.462	0.077	0.231	4.621
52	Blanchard River	24	158	0.458	0.500	0.042	0.274	7.444
53	Little Miami River	39	362	0.538	0.436	0.026	0.238	9.519

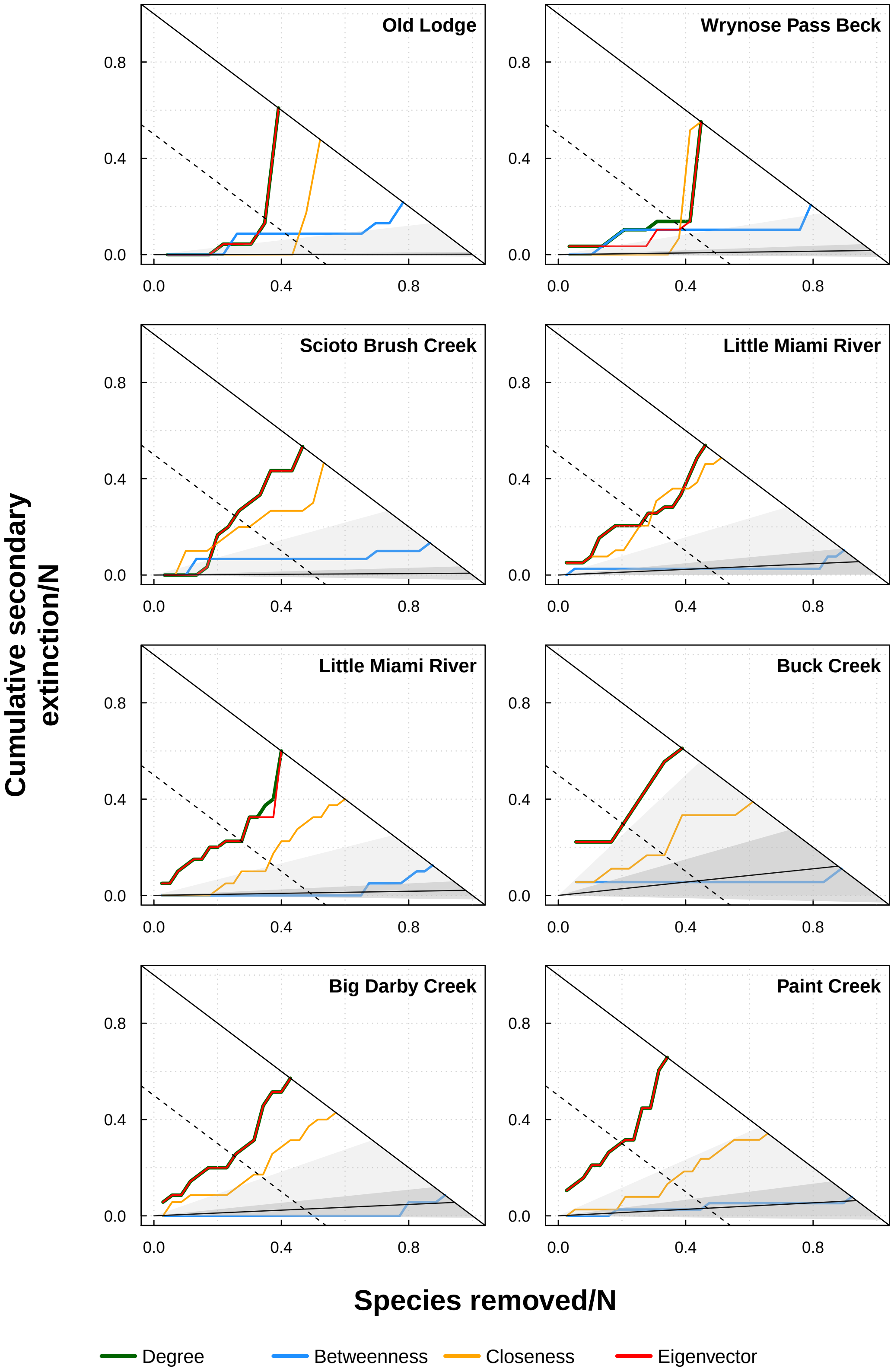
Table A2: A summary of the 53 food webs, including food web name, which ecosystems (or groups) they belong to, web size (N), link numbers (E), proportion of basal (B) / intermediate (I) / top-level species (T), connectance (C), and maximum trophic height (max TH). Food webs were ordered based on ecosystem types (1: freshwater streams; 2: lentic; 3: marine).

C.2 Food web robustness under different centrality rankings




Allt a'Mharcaidh
Allt na Coire nan Con
Beagh's Burn
Bere Stream
Broadstone
Coneyglen Burn
Dargall Lane
Duddon Beck
0.0
0.4
0.8
Cumulative secondary extinction/N
Species removed/N
Degree
Betweenness
Closeness
Eigenvector


Duddon main channel
Duddon Pike Beck
Etherow
Hardknott Gill
Lone Oak
Mill Stream
Mosendale Beck
Narrator Brook
0.0
0.4
0.8
Cumulative secondary extinction/N
Species removed/N
Degree
Betweenness
Closeness
Eigenvector


Old Lodge
Wrynose Pass Beck
Scioto Brush Creek
Little Miami River
Little Miami River
Buck Creek
Big Darby Creek
Paint Creek
0.0
0.4
0.8
Cumulative secondary extinction/N
Species removed/N
Degree
Betweenness
Closeness
Eigenvector

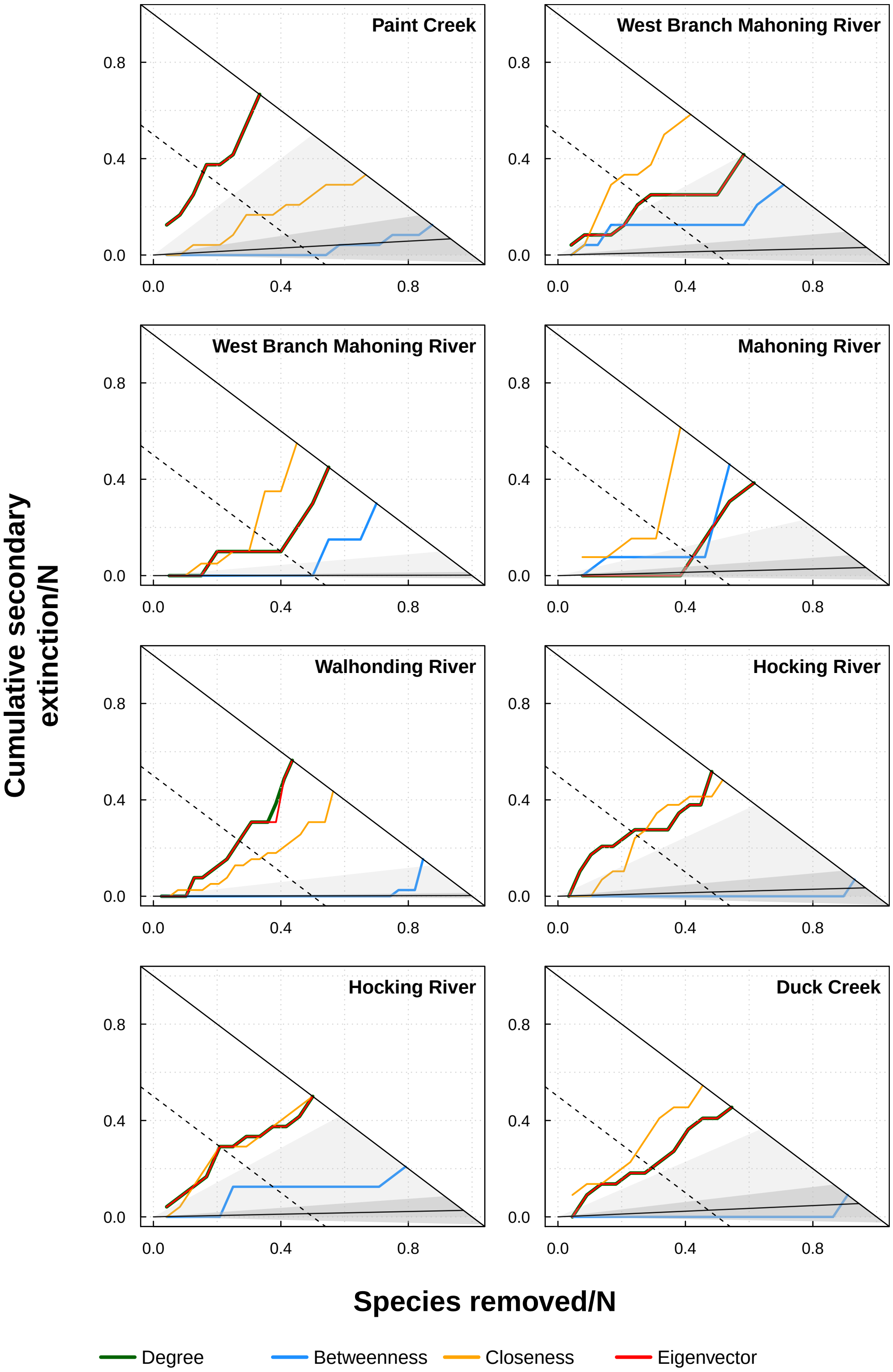


Figure A1: Robustness curve for all the food webs. Secondary extinction is shown on y-axis, which is the consequence of simulated species removal. Each removal step is displayed using a dot. Different removal strategies are shown in different colours, shown at the top-right of each plot.

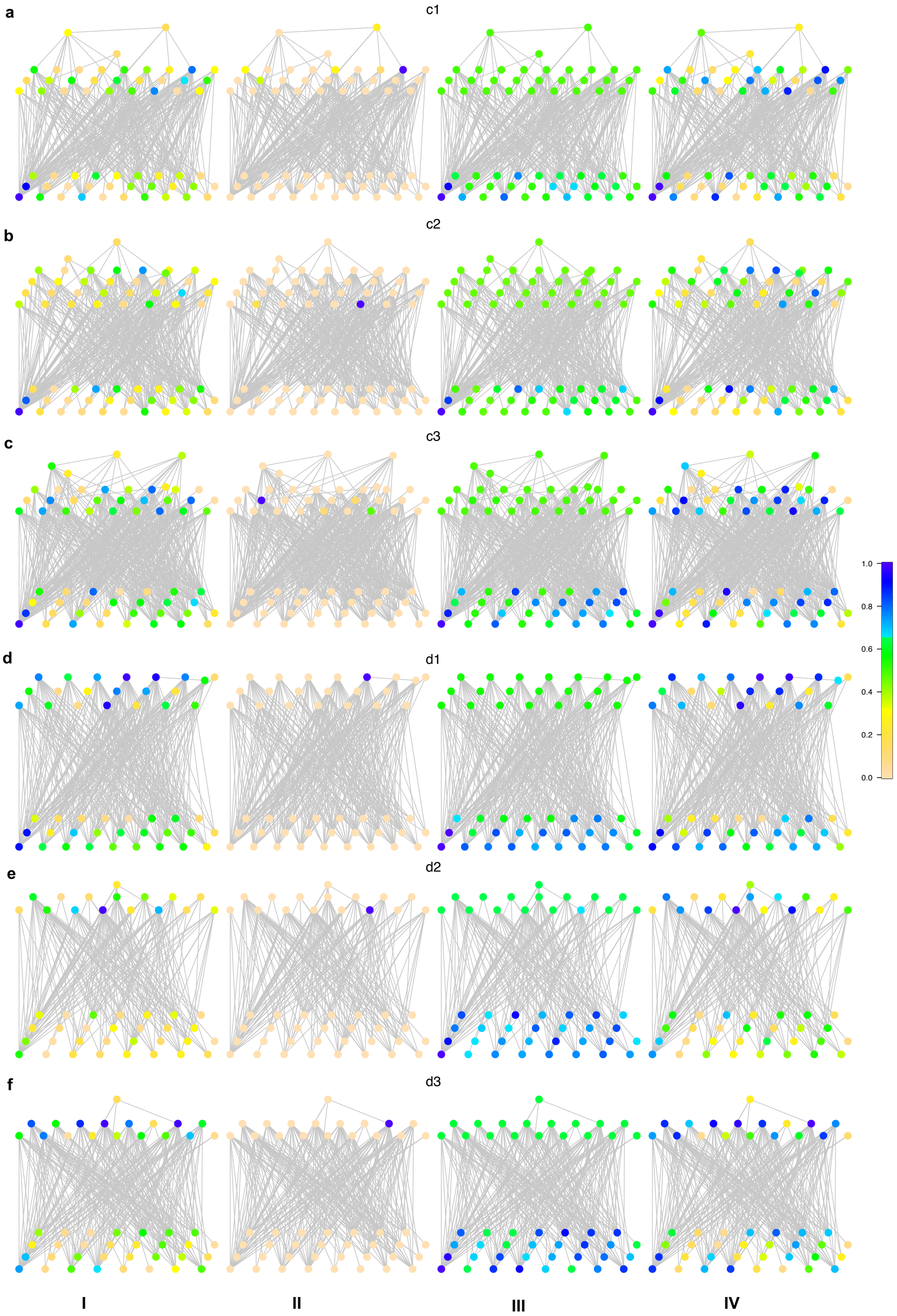
Index	Site	DC	BC	CC	EC	Random (Avg.)
1	c1	0.271	0.339	0.271	0.271	0.456
2	c2	0.254	0.286	0.270	0.286	0.468
3	c3	0.312	0.361	0.311	0.311	0.471
4	c4	0.323	0.369	0.323	0.338	0.470
5	d1	0.319	0.404	0.383	0.319	0.484
6	d2	0.152	0.413	0.370	0.196	0.461
7	d3	0.204	0.367	0.367	0.224	0.457
8	d4	0.212	0.365	0.346	0.250	0.473
9	Ythan Estuary	0.049	0.239	0.011	0.054	0.346
10	Tuesday Lake 1984	0.240	0.160	0.280	0.200	0.462
11	Tuesday Lake 1986	0.098	0.314	0.314	0.157	0.440
12	Broadstone Stream	0.185	0.500	0.407	0.222	0.469
13	Benguela	0.310	0.241	0.034	0.276	0.312
14	Skipwith Pond	0.306	0.500	0.028	0.333	0.386
15	Afon Gwy	0.375	0.417	0.375	0.417	0.500
16	Afon Hafren	0.4	0.360	0.360	0.360	0.499
17	Allt a'Mharcaidh	0.400	0.400	0.325	0.400	0.500
18	Allt na Coire nan Con	0.318	0.409	0.364	0.364	0.484
19	Beagh's Burn	0.267	0.433	0.500	0.267	0.500
20	Bere Stream	0.431	0.400	0.385	0.415	0.496
21	Broadstone	0.440	0.440	0.280	0.440	0.499
22	Coneyglen Burn	0.091	0.091	0.500	0.091	0.479
23	Dargall Lane	0.286	0.286	0.381	0.286	0.495
24	Duddon Beck	0.4	0.400	0.350	0.400	0.496
25	Duddon main channel	0.210	0.211	0.474	0.211	0.480

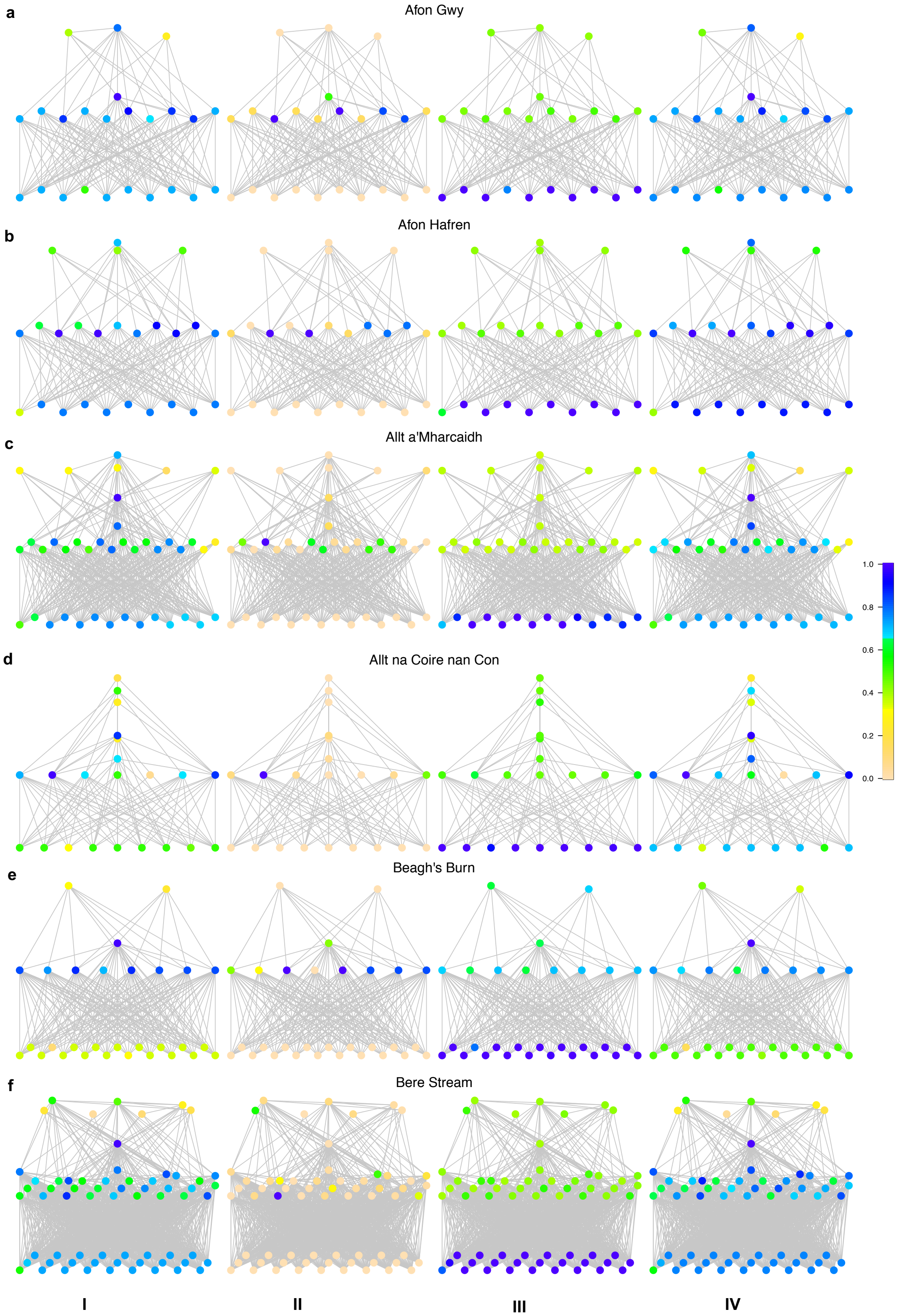
26	Duddon Pike Beck	0.371	0.400	0.343	0.429	0.499
27	Etherow	0.386	0.432	0.386	0.432	0.499
28	Hardknott Gill	0.386	0.364	0.318	0.386	0.490
29	Lone Oak	0.417	0.500	0.375	0.458	0.500
30	Mill Stream	0.425	0.402	0.368	0.402	0.497
31	Mosendale Beck	0.333	0.333	0.429	0.333	0.494
32	Narrator Brook	0.328	0.393	0.295	0.344	0.494
33	Old Lodge	0.348	0.435	0.478	0.348	0.499
34	Wrynose Pass Beck	0.379	0.414	0.414	0.414	0.491
35	Kennet	0.265	0.398	0.434	0.277	0.466
36	Scioto Brush Creek	0.233	0.467	0.333	0.267	0.496
37	Little Miami River	0.256	0.487	0.308	0.282	0.472
38	Little Miami River 2	0.275	0.500	0.375	0.300	0.489
39	Buck Creek	0.167	0.500	0.389	0.222	0.439
40	Big Darby Creek	0.257	0.500	0.343	0.257	0.472
41	Paint Creek	0.184	0.474	0.368	0.211	0.468
42	Paint Creek 2	0.125	0.500	0.375	0.167	0.466
43	West Branch Mahoning River	0.250	0.417	0.208	0.292	0.484
44	West Branch Mahoning River 2	0.400	0.500	0.350	0.450	0.499
45	Mahoning River	0.385	0.462	0.308	0.462	0.483
46	Walhonding River	0.256	0.500	0.359	0.282	0.499
47	Hocking River 2	0.208	0.500	0.276	0.241	0.483
48	Hocking River	0.250	0.417	0.250	0.250	0.487
49	Duck Creek	0.273	0.500	0.273	0.318	0.472
50	Cuyahoga River	0.333	0.333	0.250	0.417	0.486
51	Sandusky River	0.231	0.500	0.231	0.231	0.398

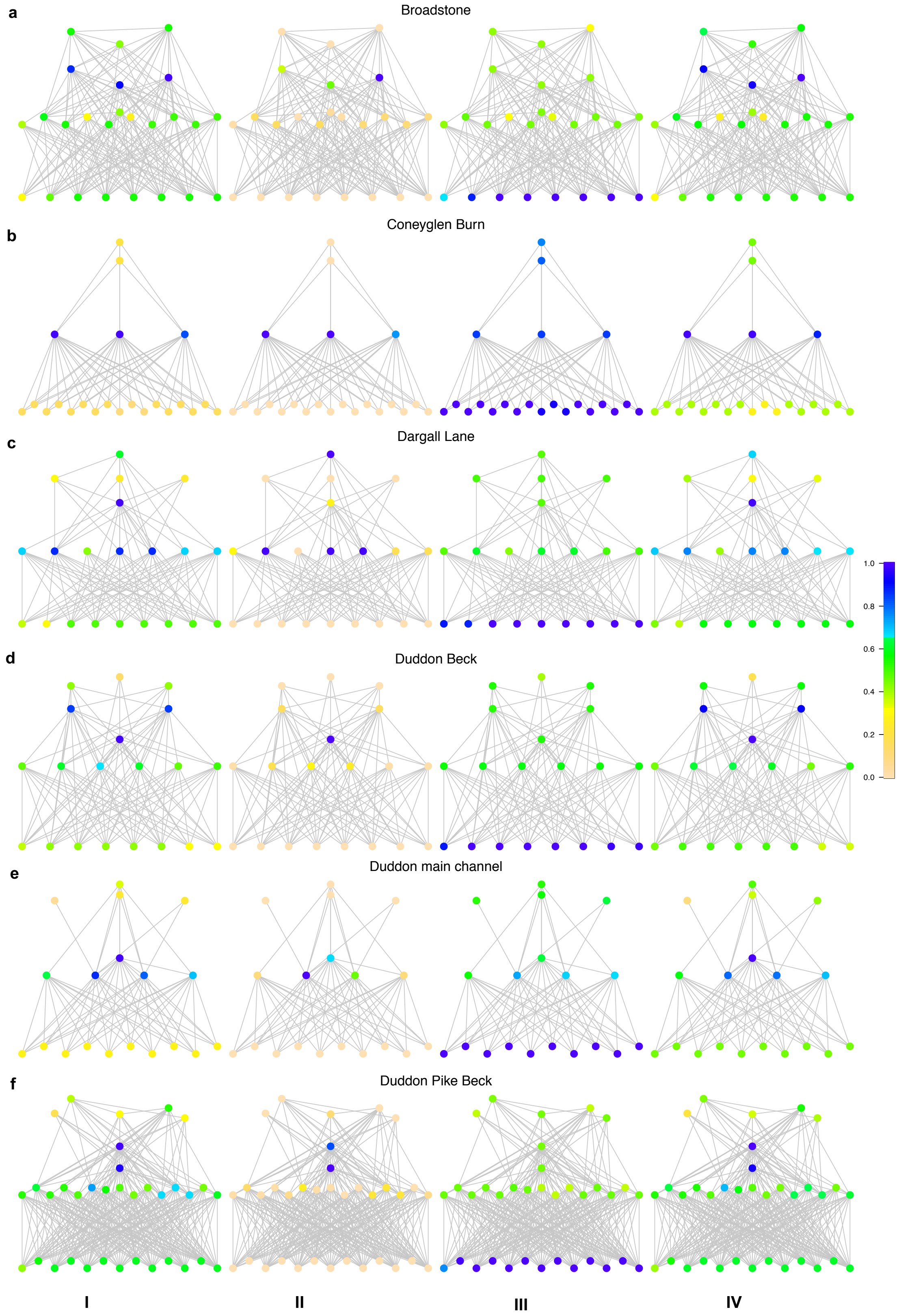
52	Blanchard River	0.292	0.458	0.292	0.333	0.467
53	Little Miami River 3	0.231	0.436	0.308	0.256	0.488

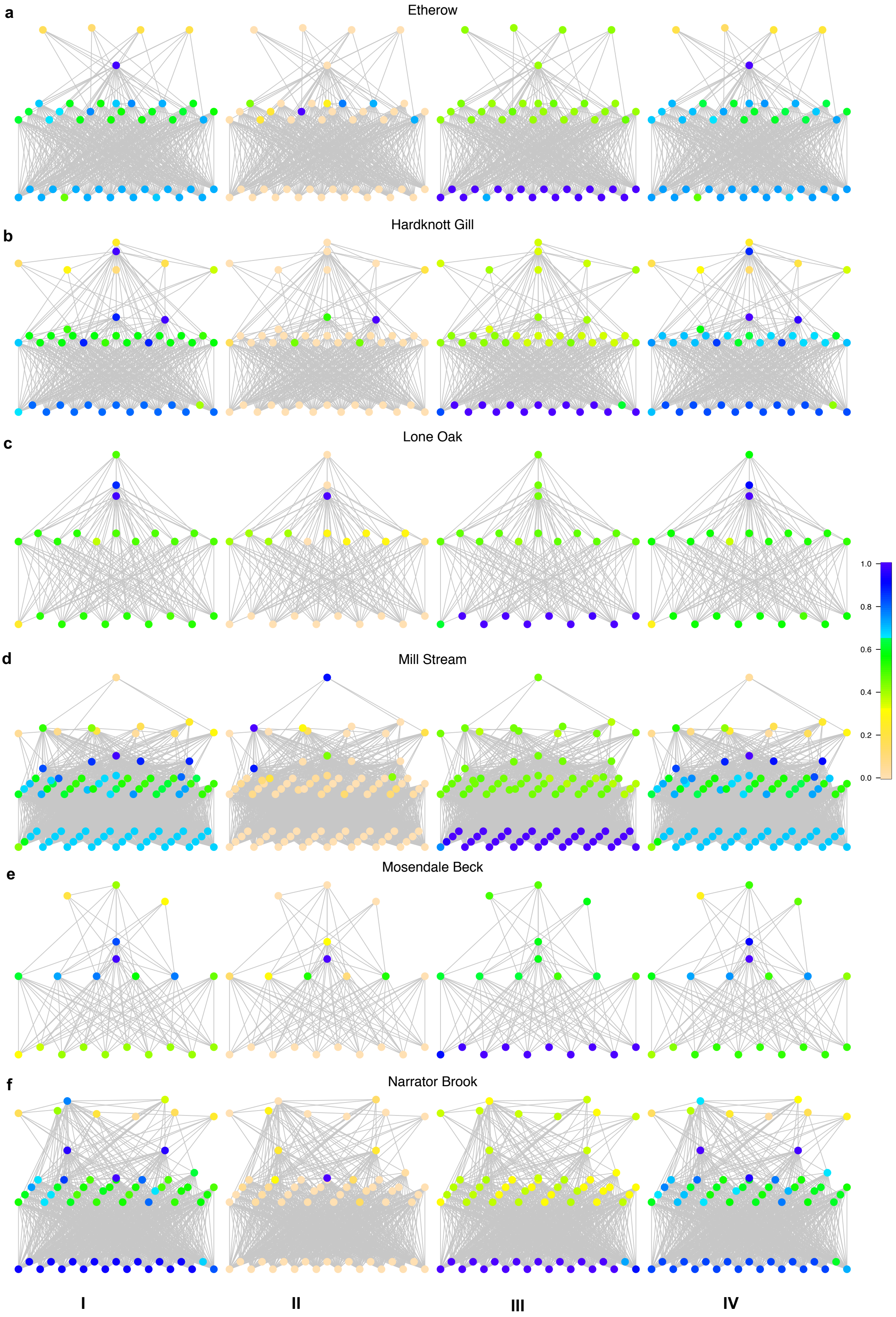
Table A3: Robustness under four centrality ranking sequences. Nodes were ordered in the decreasing order of centrality indices: Degree (DC), Betweenness (BC), Closeness (CC), and Eigenvector (EC) centrality. Results were compared with robustness under random removal.

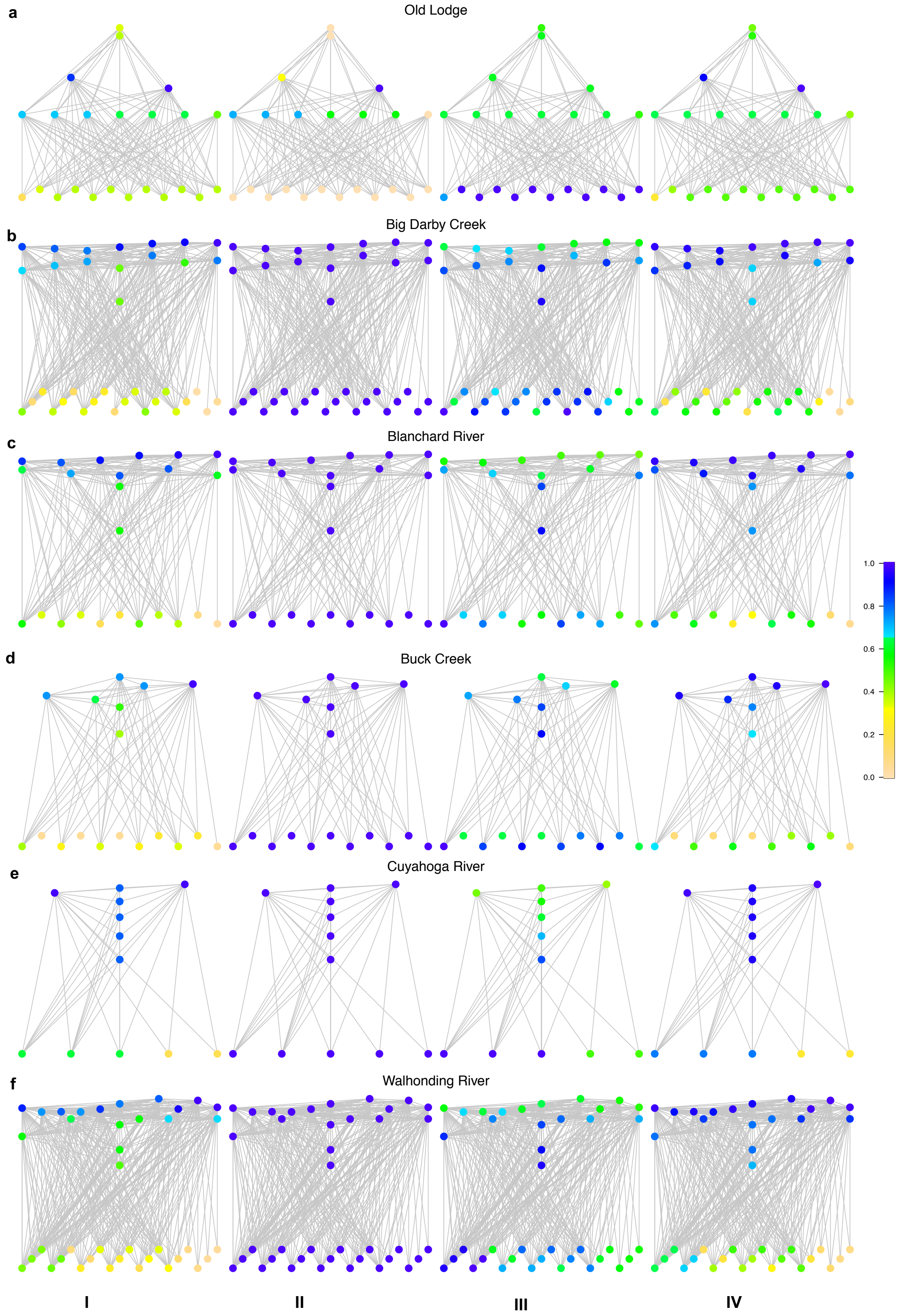
C.3 Species centrality in colour gradient

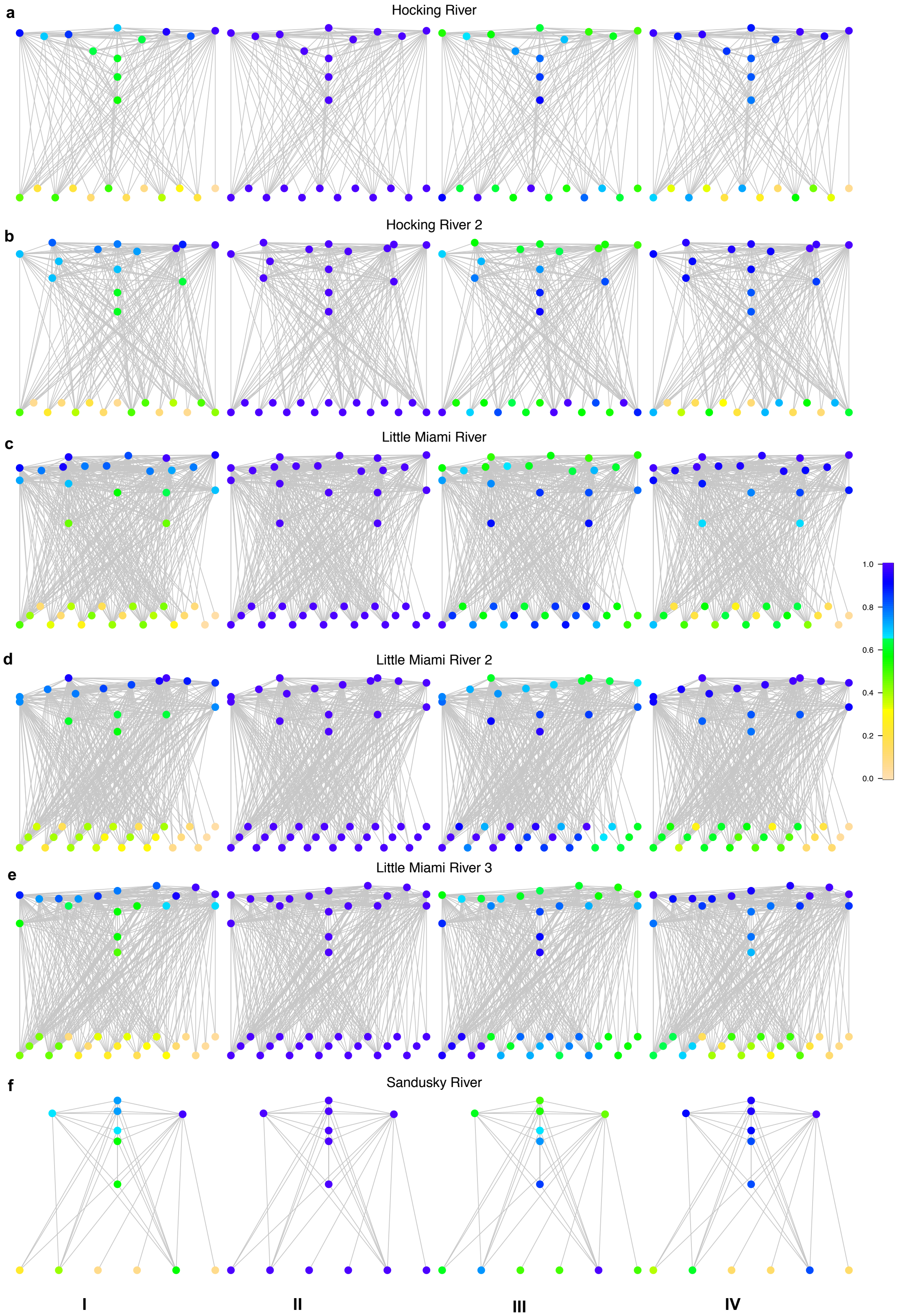


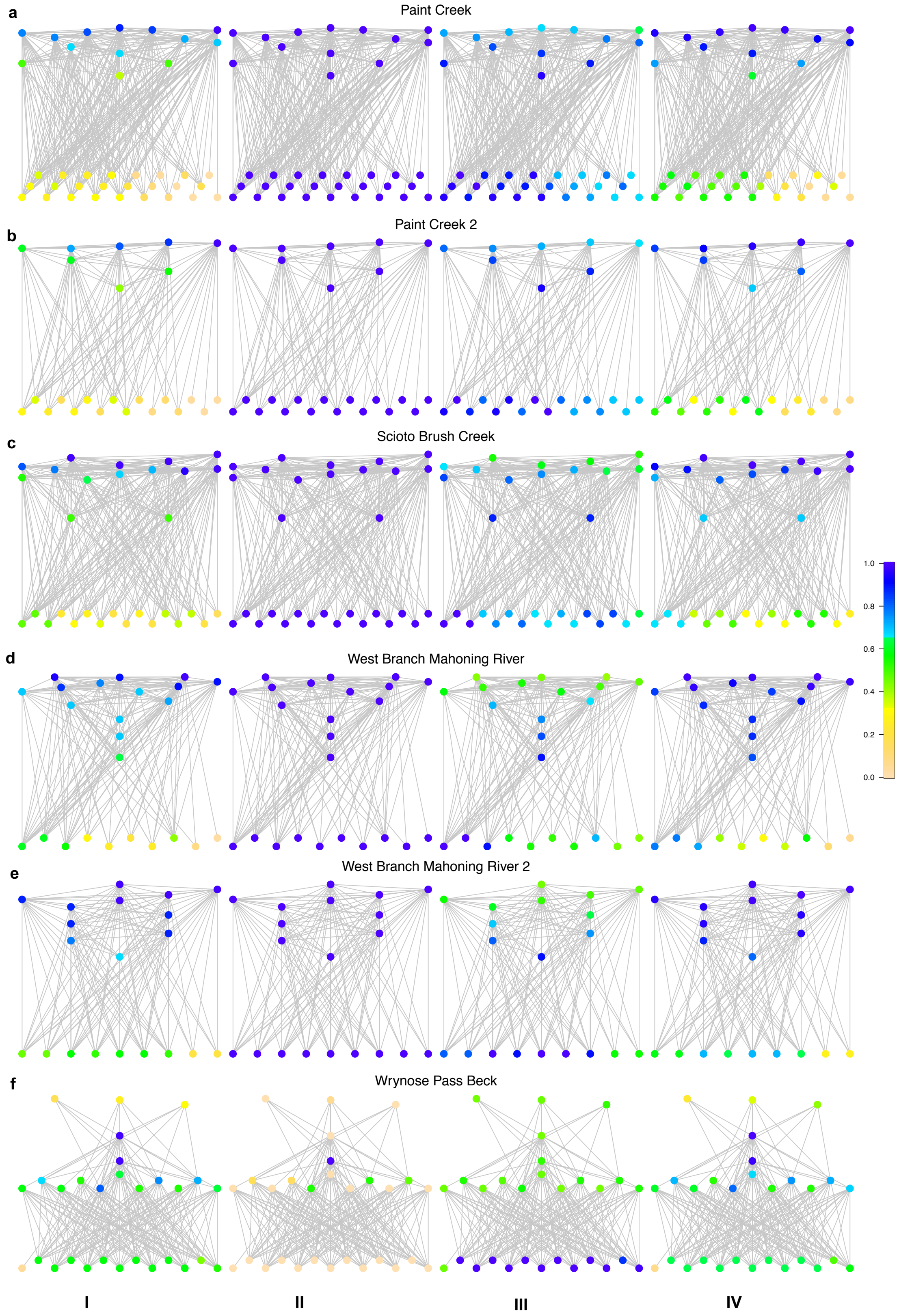

a
Afon Gwy
b
Afon Hafren
c
Allt a'Mharcaidh
d
Allt na Coire nan Con
e
Beagh's Burn
f
Bere Stream
1.0
0.8
0.6
0.4
0.2
0.0
I
II
III
IV


a
Broadstone
b
Coneyglen Burn
c
Dargall Lane
d
Duddon Beck
e
Duddon main channel
f
Duddon Pike Beck
I
II
III
IV
1.0
0.8
0.6
0.4
0.2
0.0


a
Etherow
b
Hardknott Gill
c
Lone Oak
d
Mill Stream
e
Mosendale Beck
f
Narrator Brook
1.0
0.8
0.6
0.4
0.2
0.0
I
II
III
IV


a
Old Lodge
b
Big Darby Creek
c
Blanchard River
d
Buck Creek
e
Cuyahoga River
f
Walhonding River
I
II
III
IV
1.0
0.8
0.6
0.4
0.2
0.0


a
Hocking River
b
Hocking River 2
c
Little Miami River
d
Little Miami River 2
e
Little Miami River 3
f
Sandusky River
1.0
0.8
0.6
0.4
0.2
0.0
I
II
III
IV


a
Paint Creek
b
Paint Creek 2
c
Scioto Brush Creek
d
West Branch Mahoning River
e
West Branch Mahoning River 2
f
Wrynose Pass Beck
I
II
III
IV
1.0
0.8
0.6
0.4
0.2
0.0

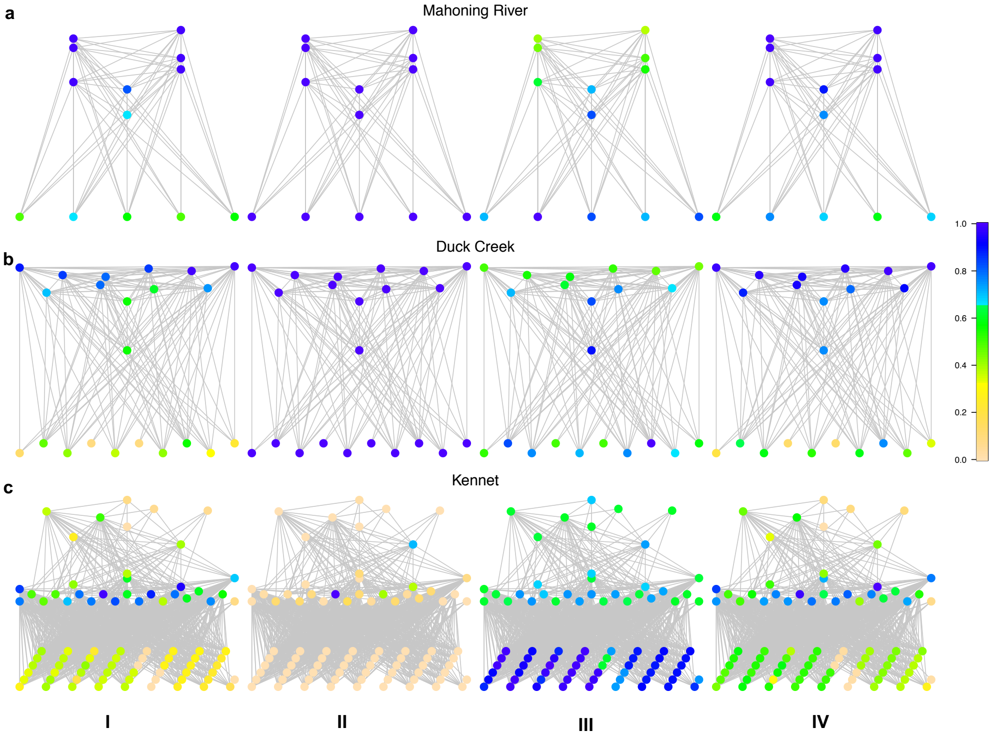
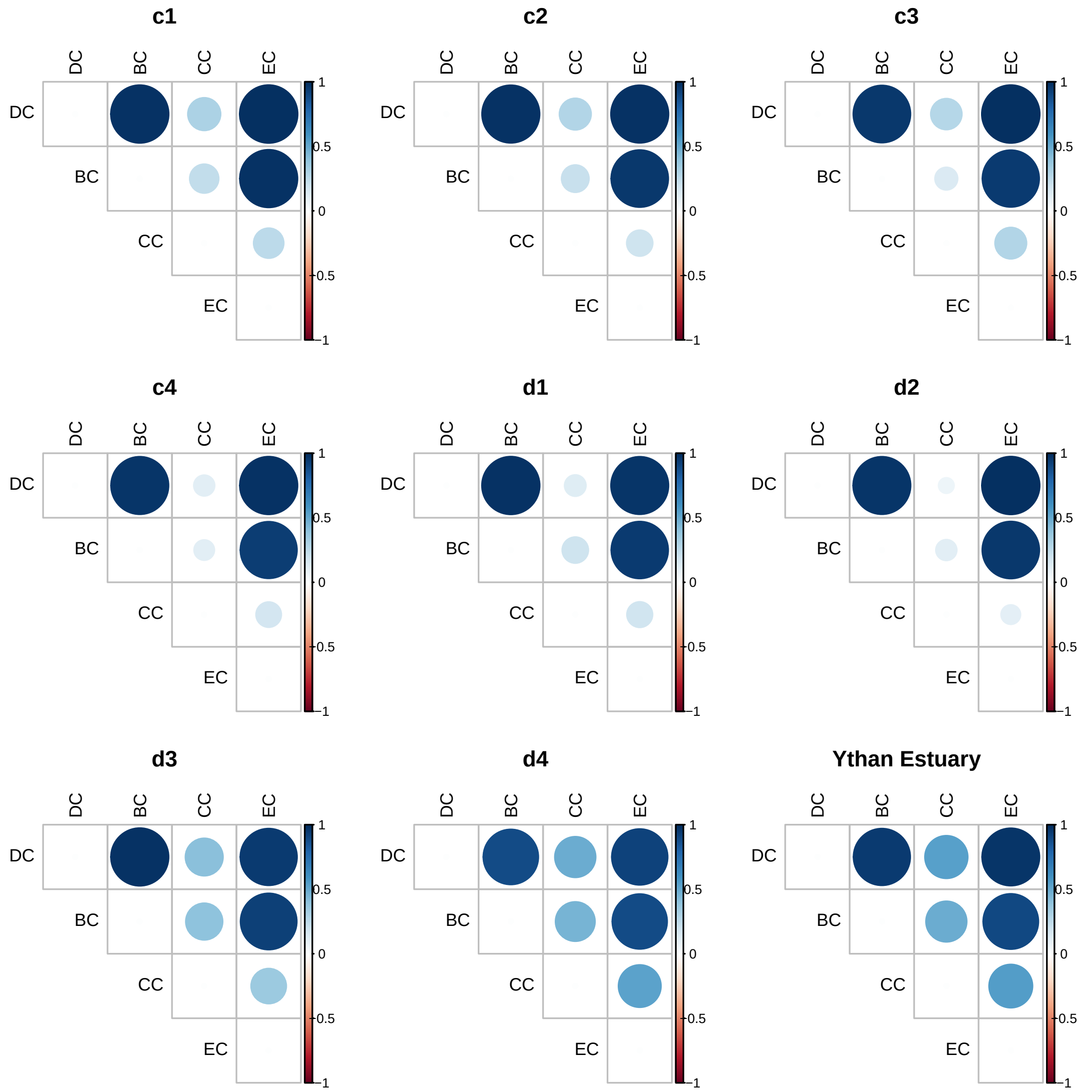
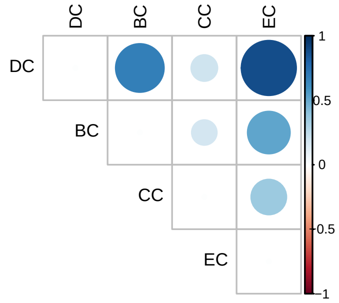
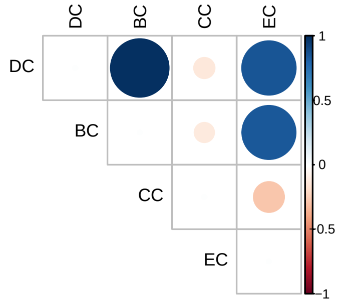


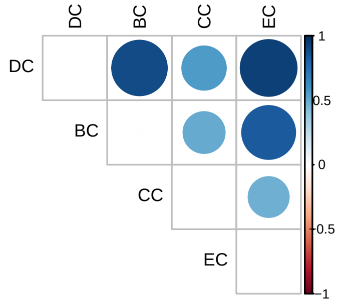
Figure A2: Centrality measurements were anlayzed on empirical food webs. Species were coloured based on different centrality rankings, including **I** - Degree centrality, **II** - Betweenness centrality, **III** - Closeness centrality, **IV** - Eigenvector centrality.

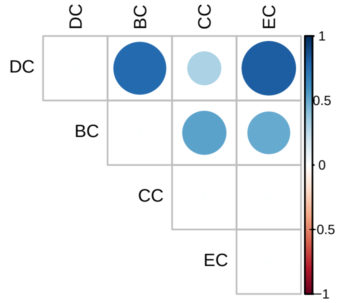
C.4 Correlation analysis among node centrality rankings

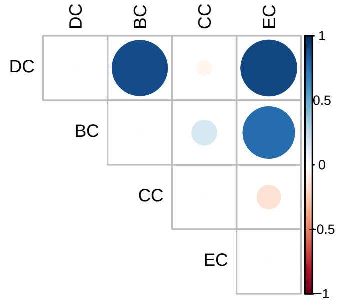


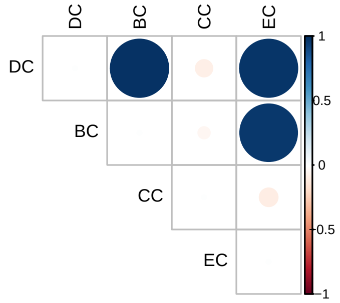

Tuesday Lake 1984

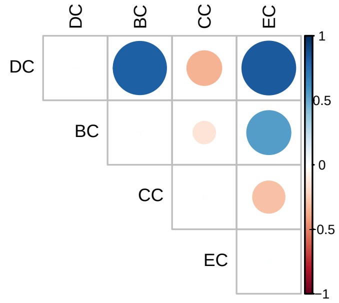

Tuesday Lake 1986

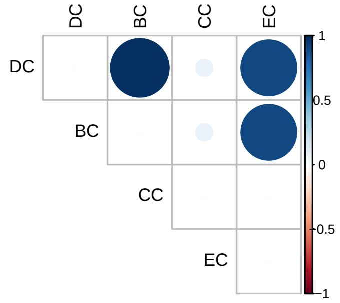

Broadstone Stream

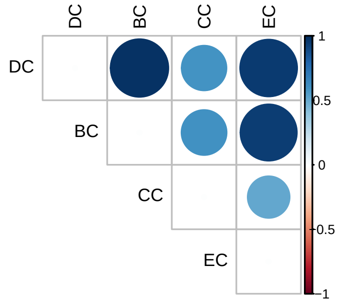

Benguela

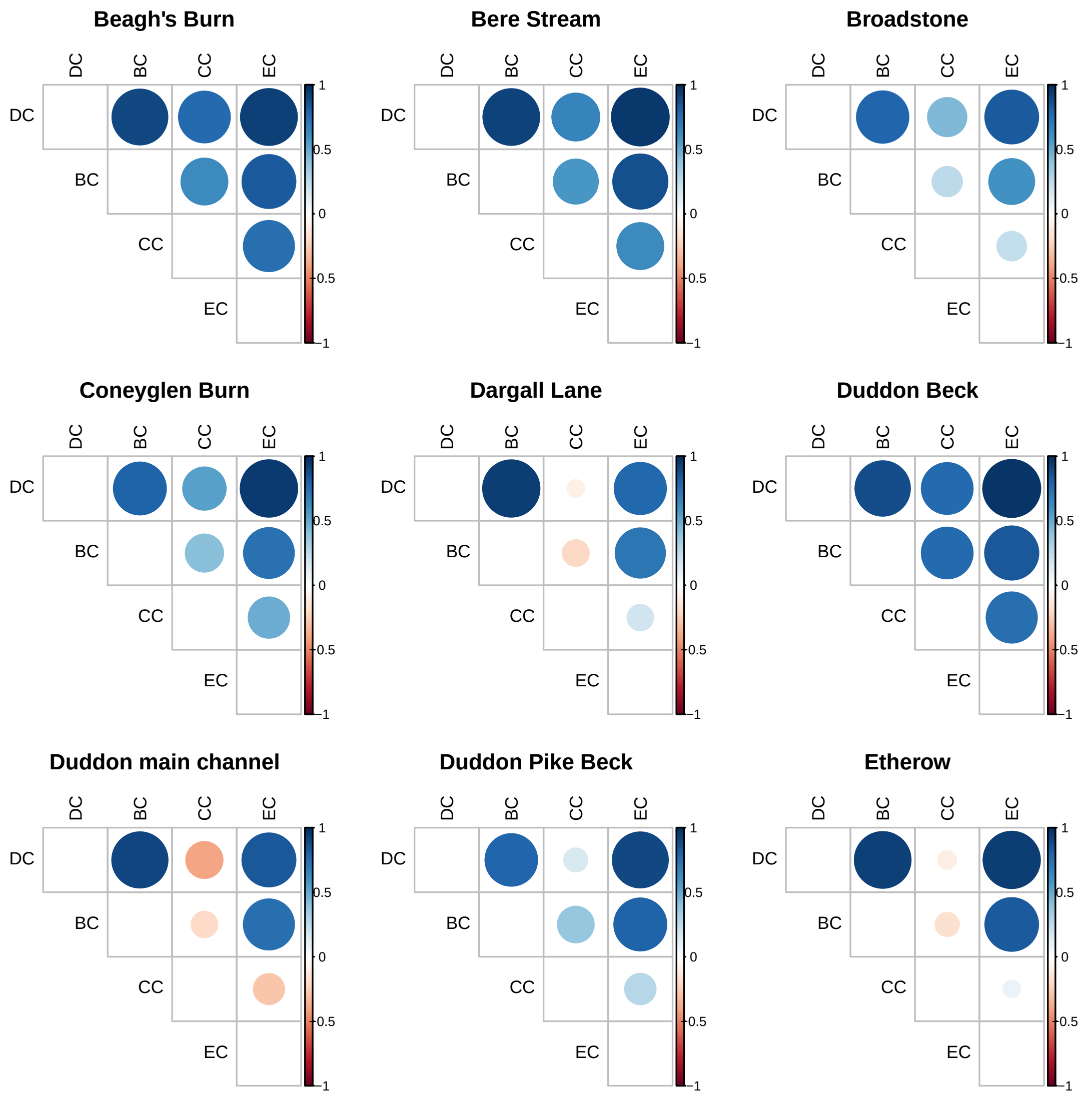

Skipwith Pond

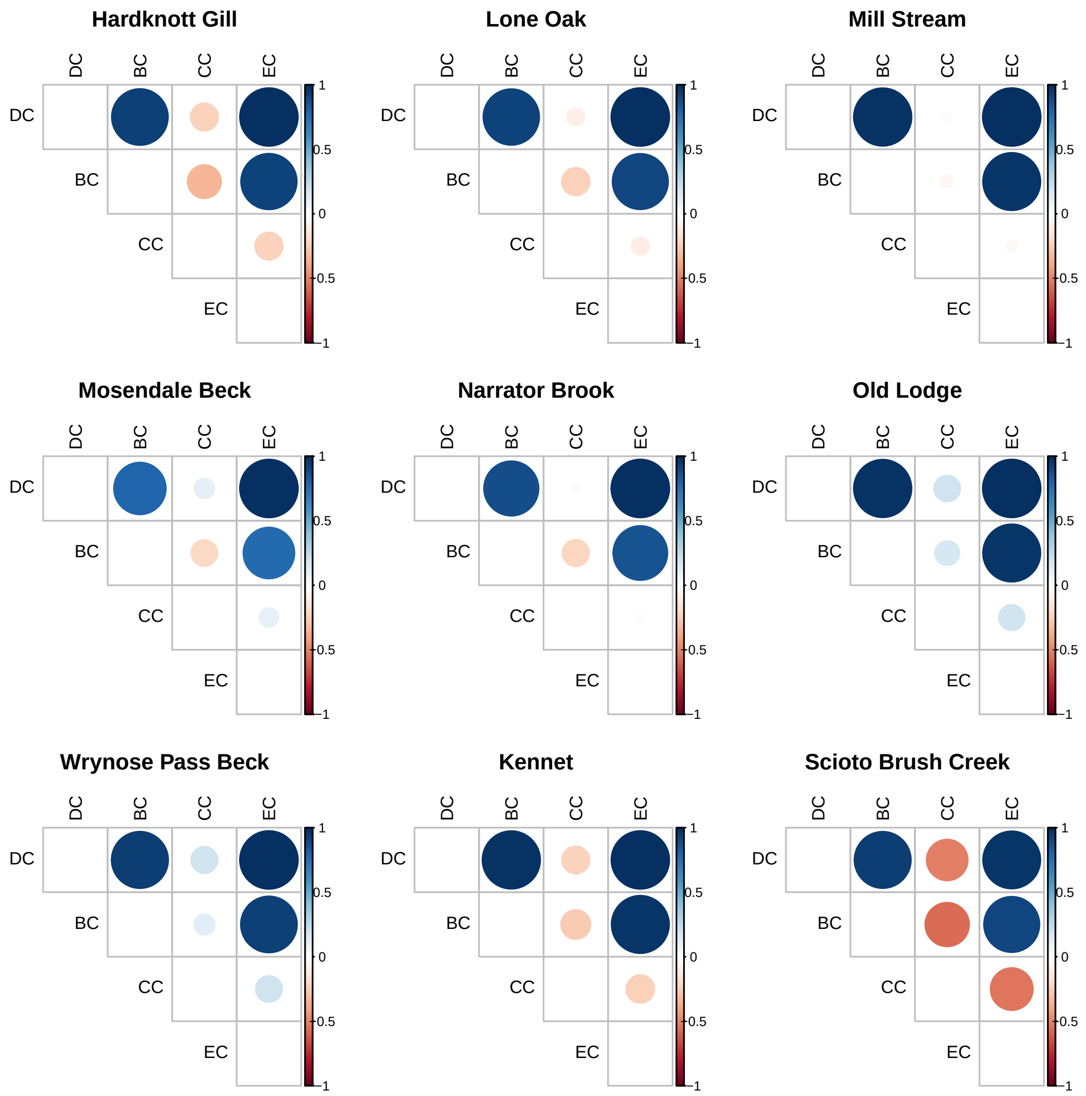

Afon Gwy

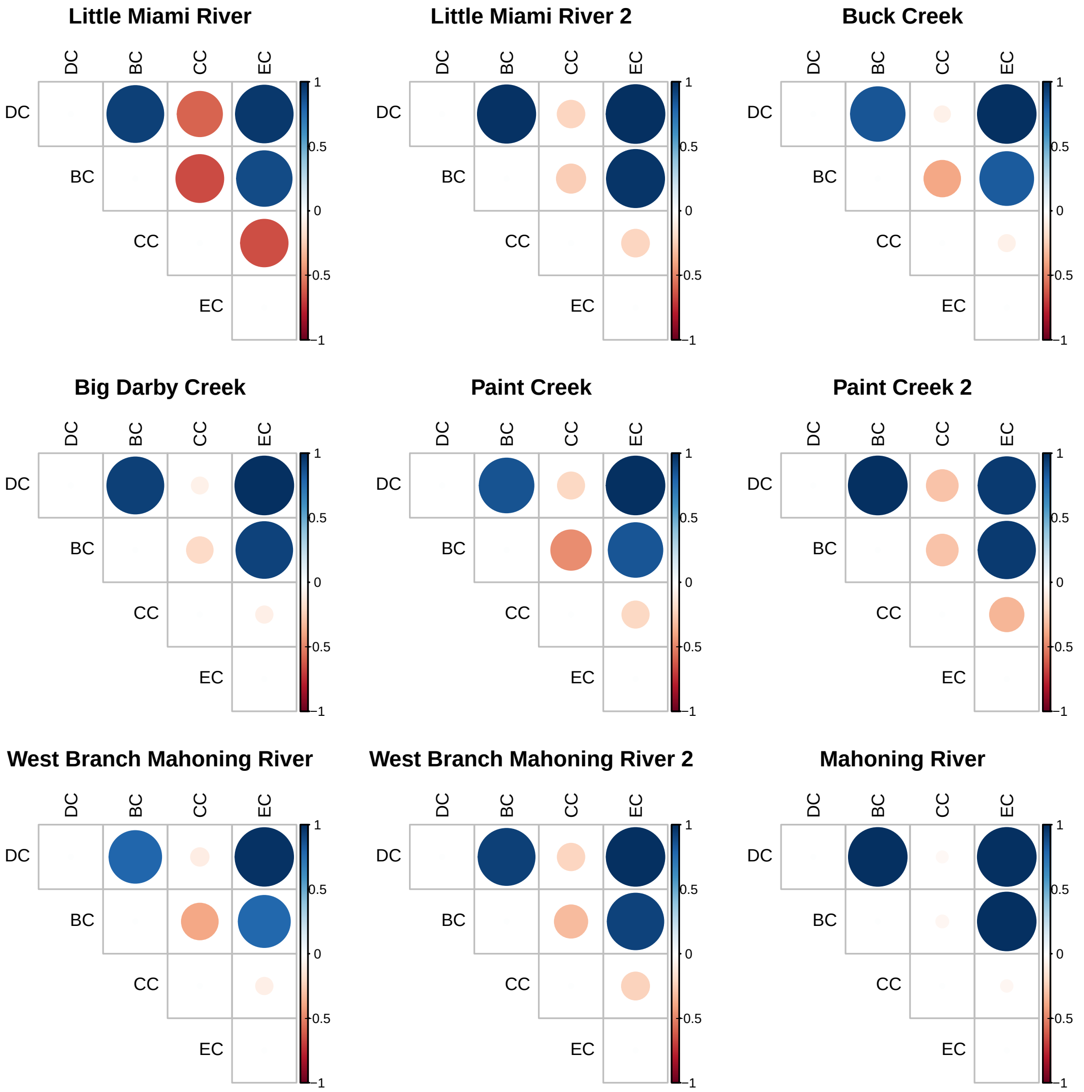

Afon Hafren


Allt a'Mharcaidh


Allt na Coire nan Con


Beagh's Burn
Bere Stream
Broadstone
Coneyglen Burn
Dargall Lane
Duddon Beck
Duddon main channel
Duddon Pike Beck
Etherow
DC
BC
CC
EC
1
0.5
0
−0.5
−1


Hardknott Gill
Lone Oak
Mill Stream
Mosendale Beck
Narrator Brook
Old Lodge
Wrynose Pass Beck
Kennet
Scioto Brush Creek
DC
BC
CC
EC
1
0.5
0
−0.5
−1



Little Miami River

Little Miami River 2

Buck Creek

Big Darby Creek

Paint Creek

Paint Creek 2

West Branch Mahoning River

West Branch Mahoning River 2

Mahoning River

Figure A3: Comparisons of the correlation between centrality metrics using the Spearman correlation coefficient. Colour and size denotes the degree of covariation with larger blue ($\rho = 1$) as the most similar and small red ($\rho = 0$) as the least similar.

References

[1] Gray, C., Baird, D.J., Baumgartner, S., Jacob, U., Jenkins, G.B., O'Gorman, E.J., Lu, X., Ma, A., Pocock, M.J., Schuwirth, N., et al.: Forum: Ecological networks: the missing links in biomonitoring science. Journal of Applied Ecology **51**(5) (2014) 1444–1449

[2] Hegland, S.J., Dunne, J., Nielsen, A., Memmott, J.: How to monitor ecological communities cost-efficiently: The example of plant–pollinator networks. Biological Conservation **143**(9) (2010) 2092–2101

[3] Mills, L., Soulé, M., Doak, D.: The keystone-species concept in ecology and conservation. BioScience **43**(4) (1993) 219–224

[4] Tylianakis, J.M., Laliberté, E., Nielsen, A., Bascompte, J.: Conservation of species interaction networks. Biological conservation **143**(10) (2010) 2270–2279

[5] Curtsdotter, A., Binzer, A., Brose, U., de Castro, F., Ebenman, B., Eklöf, A., Riede, J.O., Thierry, A., Rall, B.C.: Robustness to secondary extinctions: comparing trait-based sequential deletions in static and dynamic food webs. Basic and Applied Ecology **12**(7) (2011) 571–580

[6] Camacho, J., Guimerà, R., Amaral, L.A.N.: Robust patterns in food web structure. Physical Review Letters **88**(22) (2002) 228102

[7] Yeakel, J.D., Guimarães, P.R., Novak, M., Fox-Dobbs, K., Koch, P.L.: Probabilistic patterns of interaction: the effects of link-strength variability on food web structure. Journal of The Royal Society Interface **9**(77) (2012) 3219–3228

[8] Rolls, R.J., Leigh, C., Sheldon, F.: Mechanistic effects of low-flow hydrology on riverine ecosystems: ecological principles and consequences of alteration. Freshwater Science **31**(4) (2012) 1163–1186

[9] Rawcliffe, R., Sayer, C.D., Woodward, G., Grey, J., Davidson, T.A., IWAN JONES, J.: Back to the future: using palaeolimnology to infer long-term changes in shallow lake food webs. Freshwater Biology **55**(3) (2010) 600–613

[10] Scheffer, M., Carpenter, S.R.: Catastrophic regime shifts in ecosystems: linking theory to observation. Trends in ecology & evolution **18**(12) (2003) 648–656

[11] Allesina, S., Bodini, A.: Who dominates whom in the ecosystem? energy flow bottlenecks and cascading extinctions. Journal of Theoretical Biology **230**(3) (2004) 351–358

[12] Garlaschelli, D., Caldarelli, G., Pietronero, L.: Universal scaling relations in food webs. Nature **423**(6936) (2003) 165–168

[13] Montoya, J.M., Solé, R.V.: Topological properties of food webs: from real data to community assembly models. Oikos **102**(3) (2003) 614–622

[14] Lurgi, M., López, B.C., Montoya, J.M.: Novel communities from climate change. Philosophical Transactions of the Royal Society of London B: Biological Sciences **367**(1605) (2012) 2913–2922

[15] Melián, C.J., Bascompte, J.: Food web cohesion. Ecology **85**(2) (2004) 352–358

[16] Staniczenko, P., Lewis, O.T., Jones, N.S., Reed-Tsochas, F.: Structural dynamics and robustness of food webs. Ecology letters **13**(7) (2010) 891–899

[17] Dunne, J.A., Williams, R.J., Martinez, N.D.: Network structure and biodiversity loss in food webs: robustness increases with connectance. Ecology letters **5**(4) (2002) 558–567

[18] Ebenman, B., Jonsson, T.: Using community viability analysis to identify fragile systems and keystone species. Trends in Ecology & Evolution **20**(10) (2005) 568–575

[19] Estrada, E.: Characterization of topological keystone species: local, global and meso-scale centralities in food webs. Ecological Complexity **4**(1) (2007) 48–57

[20] Jordán, F.: Keystone species and food webs. Philosophical Transactions of the Royal Society B: Biological Sciences **364**(1524) (2009) 1733–1741

[21] Sanders, D., van Veen, F.F.: Indirect commensalism promotes persistence of secondary consumer species. Biology letters (2012) rsbl20120572

[22] Paine, R.: Food web complexity and species diversity. American Naturalist (1966) 65–75

[23] Hall, S., Raffaelli, D.: Food-web patterns: lessons from a species-rich web. The Journal of Animal Ecology (1991) 823–841

[24] Ravasz, E., Somera, A.L., Mongru, D.A., Oltvai, Z.N., Barabási, A.L.: Hierarchical

organization of modularity in metabolic networks. science **297**(5586) (2002) 1551–1555

[25] Eriksen, K.A., Simonsen, I., Maslov, S., Sneppen, K.: Modularity and extreme edges of the internet. Physical review letters **90**(14) (2003) 148701

[26] Guimera, R., Danon, L., Diaz-Guilera, A., Giralt, F., Arenas, A.: Self-similar community structure in a network of human interactions. Physical review E **68**(6) (2003) 065103

[27] Newman, M.E.: Modularity and community structure in networks. Proceedings of the National Academy of Sciences **103**(23) (2006) 8577–8582

[28] Karrer, B., Levina, E., Newman, M.E.: Robustness of community structure in networks. Physical Review E **77**(4) (2008) 046119

[29] McNaughton, S.J.: Diversity and stability of ecological communities: a comment on the role of empiricism in ecology. American Naturalist (1977) 515–525

[30] Cabral, J.S., Kreft, H.: Linking ecological niche, community ecology and biogeography: insights from a mechanistic niche model. Journal of Biogeography **39**(12) (2012) 2212–2224

[31] MacArthur, R.: Fluctuations of animal populations and a measure of community stability. ecology **36**(3) (1955) 533–536

[32] Cohen, J.E., Jonsson, T., Carpenter, S.R.: Ecological community description using the food web, species abundance, and body size. Proceedings of the National Academy of Sciences **100**(4) (2003) 1781–1786

[33] McCann, K., Hastings, A.: Re–evaluating the omnivory–stability relationship in food webs. Proceedings of the Royal Society of London. Series B: Biological Sciences **264**(1385) (1997) 1249–1254

[34] McCann, K., Hastings, A., Huxel, G.: Weak trophic interactions and the balance of nature. Nature **395**(6704) (1998) 794–798

[35] Ledger, M.E., Edwards, F.K., Brown, L.E., Milner, A.M., Woodward, G.: Impact of simulated drought on ecosystem biomass production: an experimental test in stream mesocosms. Global Change Biology **17**(7) (2011) 2288–2297

[36] Woodward, G., Brown, L.E., Edwards, F.K., Hudson, L.N., Milner, A.M., Reuman,

D.C., Ledger, M.E.: Climate change impacts in multispecies systems: drought alters food web size structure in a field experiment. Philosophical Transactions of the Royal Society B: Biological Sciences **367**(1605) (2012) 2990–2997

[37] Bascompte, J., Jordano, P., Melián, C.J., Olesen, J.M.: The nested assembly of plant–animal mutualistic networks. Proceedings of the National Academy of Sciences **100**(16) (2003) 9383–9387

[38] Williams, R., Martinez, N., et al.: Simple rules yield complex food webs. Nature **404**(6774) (2000) 180–183

[39] Williams, R., Martinez, N.: Success and its limits among structural models of complex food webs. Journal of Animal Ecology **77**(3) (2008) 512–519

[40] Dunne, J.A., Williams, R.J., Martinez, N.D.: Food-web structure and network theory: the role of connectance and size. Proceedings of the National Academy of Sciences **99**(20) (2002) 12917–12922

[41] Bersier, L.F., Banaek-Richter, C., Cattin, M.F.: Quantitative descriptors of food-web matrices. Ecology **83**(9) (2002) 2394–2407

[42] Estrada, E.: Food webs robustness to biodiversity loss: the roles of connectance, expansibility and degree distribution. Journal of theoretical biology **244**(2) (2007) 296–307

[43] Gilbert, A.J.: Connectance indicates the robustness of food webs when subjected to species loss. Ecological Indicators **9**(1) (2009) 72–80

[44] Martinez, N.: Artifacts or attributes? effects of resolution on the little rock lake food web. Ecological Monographs (1991) 367–392

[45] Pawar, S., Dell, A., Savage, V.: Dimensionality of consumer search space drives trophic interaction strengths. Nature **486**(7404) (2012) 485–489

[46] Johnson, S., Domínguez-García, V., Donetti, L., Muñoz, M.A.: Trophic coherence determines food-web stability. Proceedings of the National Academy of Sciences **111**(50) (2014) 17923–17928

[47] Kaunzinger, C.M., Morin, P.J.: Productivity controls food-chain properties in microbial communities. Nature **395**(6701) (1998) 495–497

[48] Mendlewicz, J., RAINER, J.D.: Number of trophic levels in ecological communi-

ties. Nature **268** (1977) 28

[49] Woodward, G., Dybkjaer, J.B., Ólafsson, J.S., Gíslason, G.M., Hannesdóttir, E.R., Friberg, N.: Sentinel systems on the razor's edge: effects of warming on arctic geothermal stream ecosystems. Global Change Biology **16**(7) (2010) 1979–1991

[50] Ledger, M.E., Brown, L.E., Edwards, F.K., Milner, A.M., Woodward, G.: Drought alters the structure and functioning of complex food webs. Nature Climate Change **3**(3) (2013) 223–227

[51] Krause, A.E., Frank, K.A., Mason, D.M., Ulanowicz, R.E., Taylor, W.W.: Compartments revealed in food-web structure. Nature **426**(6964) (2003) 282–285

[52] Da Silva, M.R., Ma, H., Zeng, A.P.: Centrality, network capacity, and modularity as parameters to analyze the core-periphery structure in metabolic networks. Proceedings of the IEEE **96**(8) (2008) 1411–1420

[53] Abbasi, A., Hossain, L., Uddin, S., Rasmussen, K.J.: Evolutionary dynamics of scientific collaboration networks: multi-levels and cross-time analysis. Scientometrics **89**(2) (2011) 687–710

[54] Csermely, P., London, A., Wu, L.Y., Uzzi, B.: Structure and dynamics of core/periphery networks. Journal of Complex Networks **1**(2) (2013) 93–123

[55] Csete, M., Doyle, J.: Bow ties, metabolism and disease. TRENDS in Biotechnology **22**(9) (2004) 446–450

[56] Liu, Y.Y., Slotine, J.J., Barabási, A.L.: Controllability of complex networks. Nature **473**(7346) (2011) 167–173

[57] Rombach, M.P., Porter, M.A., Fowler, J.H., Mucha, P.J.: Core-periphery structure in networks. SIAM Journal on Applied mathematics **74**(1) (2014) 167–190

[58] Bassett, D.S., Wymbs, N.F., Rombach, M.P., Porter, M.A., Mucha, P.J., Grafton, S.T.: Task-based core-periphery organization of human brain dynamics. PLoS computational biology **9**(9) (2013) e1003171

[59] Yang, Y., Li, Z., Chen, Y., Zhang, X., Wang, S.: Improving the robustness of complex networks with preserving community structure. PloS one **10**(2) (2015) e0116551

[60] Dunne, J.: The network structure of food webs. Ecological networks: linking

structure to dynamics in food webs (2006) 27–86

[61] Girvan, M., Newman, M.E.: Community structure in social and biological networks. Proceedings of the National Academy of Sciences **99**(12) (2002) 7821–7826

[62] Newman, M.E.: The structure and function of complex networks. SIAM review **45**(2) (2003) 167–256

[63] Newman, M.E.: Detecting community structure in networks. The European Physical Journal B-Condensed Matter and Complex Systems **38**(2) (2004) 321–330

[64] Holme, P.: Core-periphery organization of complex networks. Physical Review E **72**(4) (2005) 046111

[65] Albert, R., Barabási, A.L.: Dynamics of complex systems: Scaling laws for the period of boolean networks. Physical Review Letters **84**(24) (2000) 5660

[66] Cohen, R., Erez, K., Ben-Avraham, D., Havlin, S.: Resilience of the internet to random breakdowns. Physical review letters **85**(21) (2000) 4626

[67] Callaway, D.S., Newman, M.E., Strogatz, S.H., Watts, D.J.: Network robustness and fragility: Percolation on random graphs. Physical review letters **85**(25) (2000) 5468

[68] Cohen, R., Erez, K., Ben-Avraham, D., Havlin, S.: Breakdown of the internet under intentional attack. Physical Review Letters **86**(16) (2001) 3682

[69] Keeling, M.J., Eames, K.T.: Networks and epidemic models. Journal of the Royal Society Interface **2**(4) (2005) 295–307

[70] Smart, A.G., Amaral, L.A., Ottino, J.M.: Cascading failure and robustness in metabolic networks. Proceedings of the National Academy of Sciences **105**(36) (2008) 13223–13228

[71] Bellingeri, M., Cassi, D., Vincenzi, S.: Efficiency of attack strategies on complex model and real-world networks. Physica A: Statistical Mechanics and its Applications **414** (2014) 174–180

[72] Van Mieghem, P., Doerr, C., Wang, H., Hernandez, J.M., Hutchison, D., Karaliopoulos, M., Kooij, R.: A framework for computing topological network robustness. Delft University of Technology, Report20101218 (2010)

[73] Hasegawa, T., Konno, K., Nemoto, K.: Robustness of correlated networks against propagating attacks. The European Physical Journal B **85**(8) (2012) 1–6

[74] Kitching, R.: Food webs in phytotelmata:?bottom-up? and ?top-down? explanations for community structure. Annual review of entomology **46**(1) (2001) 729–760

[75] Eklöf, A., Ebenman, B.: Species loss and secondary extinctions in simple and complex model communities. Journal of animal ecology **75**(1) (2006) 239–246

[76] Binzer, A., Brose, U., Curtsdotter, A., Eklöf, A., Rall, B.C., Riede, J.O., de Castro, F.: The susceptibility of species to extinctions in model communities. Basic and Applied Ecology **12**(7) (2011) 590–599

[77] Thierry, A., Beckerman, A.P., Warren, P.H., Williams, R.J., Cole, A.J., Petchey, O.L.: Adaptive foraging and the rewiring of size-structured food webs following extinctions. Basic and Applied Ecology **12**(7) (2011) 562–570

[78] Kondoh, M.: Foraging adaptation and the relationship between food-web complexity and stability. Science **299**(5611) (2003) 1388–1391

[79] C Rall, B., Guill, C., Brose, U.: Food-web connectance and predator interference dampen the paradox of enrichment. Oikos **117**(2) (2008) 202–213

[80] Dunne, J.A., Williams, R.J., Martinez, N.D.: Network structure and robustness of marine food webs. Marine Ecology Progress Series **273** (2004) 291–302

[81] Layer, K., Riede, J.O., Hildrew, A.G., Woodward, G.: Food web structure and stability in 20 streams across a wide ph gradient. Advances In Ecological Research **42** (2010) 265–299

[82] Memmott, J., Waser, N.M., Price, M.V.: Tolerance of pollination networks to species extinctions. Proceedings of the Royal Society of London. Series B: Biological Sciences **271**(1557) (2004) 2605–2611

[83] Dunne, J.A., Williams, R.J.: Cascading extinctions and community collapse in model food webs. Philosophical Transactions of the Royal Society B: Biological Sciences **364**(1524) (2009) 1711–1723

[84] Pocock, M.J., Evans, D.M., Memmott, J.: The robustness and restoration of a network of ecological networks. Science **335**(6071) (2012) 973–977

[85] Jordán, F., Liu, W.c., Davis, A.J.: Topological keystone species: measures of

positional importance in food webs. Oikos **112**(3) (2006) 535–546

[86] Jordán, F., Liu, W.c., Mike, Á.: Trophic field overlap: a new approach to quantify keystone species. Ecological Modelling **220**(21) (2009) 2899–2907

[87] Allesina, S., Pascual, M.: Googling food webs: can an eigenvector measure species' importance for coextinctions? PLoS computational biology **5**(9) (2009) e1000494

[88] Freeman, L.C.: A set of measures of centrality based on betweenness. Sociometry (1977) 35–41

[89] Joy, M., Brock, A., Ingber, D., Huang, S.: High-betweenness proteins in the yeast protein interaction network. Journal of Biomedicine and Biotechnology **2005**(2) (2005) 96–103

[90] Lu, X., Gray, C., Brown, L.E., Ledger, M.E., Milner, A.M., Mondragón, R.J., Woodward, G., Ma, A.: Drought rewires the cores of food webs. Nature Climate Change (2016)

[91] Freeman, L.C.: Centrality in social networks conceptual clarification. Social networks **1**(3) (1979) 215–239

[92] Molloy, M., Reed, B.: A critical point for random graphs with a given degree sequence. Random Structures & Algorithms **6**(2-3) (1995) 161–180

[93] Borgatti, S.P., Carley, K.M., Krackhardt, D.: On the robustness of centrality measures under conditions of imperfect data. Social networks **28**(2) (2006) 124–136

[94] Barabási, A., Albert, R.: Emergence of scaling in random networks. science **286**(5439) (1999) 509–512

[95] Albert, R., Jeong, H., Barabási, A.: Error and attack tolerance of complex networks. Nature **406**(6794) (2000) 378–382

[96] Edwards, J.S., Palsson, B.O.: Robustness analysis of the escherichiacoli metabolic network. Biotechnology Progress **16**(6) (2000) 927–939

[97] Latora, V., Marchiori, M.: Efficient behavior of small-world networks. Physical Review Letters **87**(19) (2001) 198701

[98] Jonsson, T., Berg, S., Pimenov, A., Palmer, C., Emmerson, M.: The reliability of r50 as a measure of vulnerability of food webs to sequential species deletions.

Oikos **124**(4) (2015) 446–457

[99] Borgatti, S.P., Everett, M.G.: Models of core/periphery structures. Social networks **21**(4) (1999) 375–395

[100] Wang, P., Lü, J., Yu, X.: Identification of important nodes in directed biological networks: A network motif approach. PLOS ONE **9**(8) (2014) e106132

[101] Della Rossa, F., Dercole, F., Piccardi, C.: Profiling core-periphery network structure by random walkers. Scientific reports **3** (2013)

[102] Ma, A., Mondragón, R.J.: Rich-cores in networks. PloS one **10**(3) (2015) e0119678

[103] Ledger, M.E., Brown, L.E., Edwards, F.K., Milner, A.M., Woodward, G.: Drought alters the structure and functioning of complex food webs. Nature Climate Change **3**(3) (2012) 223–227

[104] Tamaddoni-Nezhad, A., Milani, G.A., Raybould, A., Muggleton, S., Bohan, D.A.: Construction and validation of food webs using logic-based machine learning and text mining. Advances in Ecological Research **49** (2013) 225–289

[105] Bohan, D.A., Caron-Lormier, G., Muggleton, S., Raybould, A., Tamaddoni-Nezhad, A.: Automated discovery of food webs from ecological data using logic-based machine learning. PloS one **6**(12) (2011) e29028

[106] Tamaddoni-Nezhad, A., Bohan, D., Raybould, A., Muggleton, S.H.: Machine learning a probabilistic network of ecological interactions. In: Inductive Logic Programming. Springer (2011) 332–346

[107] Hudson, L.N., Emerson, R., Jenkins, G.B., Layer, K., Ledger, M.E., Pichler, D.E., Thompson, M.S., O'Gorman, E.J., Woodward, G., Reuman, D.C.: Cheddar: analysis and visualisation of ecological communities in R. Methods in Ecology and Evolution **4**(1) (2013) 99–104

[108] Layer, K., Hildrew, A.G., Jenkins, G.B., Riede, J.O., Rossiter, S.J., Townsend, C.R., Woodward, G.: Long-term dynamics of a well-characterised food web: four decades of acidification and recovery in the broadstone stream model system. Advances in ecological research **44** (2011) 69–117

[109] Thompson, M.S., Bankier, C., Bell, T., Dumbrell, A.J., Gray, C., Ledger, M.E., Lehmann, K., McKew, B.A., Sayer, C.D., Shelley, F., et al.: Gene-to-ecosystem

impacts of a catastrophic pesticide spill: testing a multilevel bioassessment approach in a river ecosystem. Freshwater Biology (2015)

[110] Dunne, J.A., Lafferty, K.D., Dobson, A.P., Hechinger, R.F., Kuris, A.M., Martinez, N.D., McLaughlin, J.P., Mouritsen, K.N., Poulin, R., Reise, K., et al.: Parasites affect food web structure primarily through increased diversity and complexity. PLoS Biol **11**(6) (2013) e1001579

[111] Neutel, A., Heesterbeek, J., van de Koppel, J., Hoenderboom, G., Vos, A., Kaldeway, C., Berendse, F., de Ruiter, P.: Reconciling complexity with stability in naturally assembling food webs. Nature **449**(7162) (2007) 599–602

[112] Estrada, E.: Topological structural classes of complex networks. Physical Review E **75**(1) (2007) 016103

[113] Estrada, E.: The structure of complex networks: theory and applications. Oxford University Press (2012)

[114] Bascompte, J., Jordano, P.: Plant-animal mutualistic networks: the architecture of biodiversity. Annual Review of Ecology, Evolution, and Systematics (2007) 567–593

[115] Ings, T.C., Montoya, J.M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C.F., Edwards, F., Figueroa, D., Jacob, U., Jones, J.I., et al.: Review: Ecological networks–beyond food webs. Journal of Animal Ecology **78**(1) (2009) 253–269

[116] Zhou, S., Mondragón, R.J.: The rich-club phenomenon in the internet topology. Communications Letters, IEEE **8**(3) (2004) 180–182

[117] McAuley, J., da Fontoura Costa, L., Caetano, T.: Rich-club phenomenon across complex network hierarchies. Applied Physics Letters **91**(8) (2007) 084103–084103

[118] Opsahl, T., Colizza, V., Panzarasa, P., Ramasco, J.J.: Prominence and control: The weighted rich-club effect. Physical review letters **101**(16) (2008) 168702

[119] Harriger, L., Van Den Heuvel, M.P., Sporns, O.: Rich club organization of macaque cerebral cortex and its role in network communication. PloS one **7**(9) (2012) e46497

[120] Colizza, V., Flammini, A., Serrano, M., Vespignani, A.: Detecting rich-club ordering in complex networks. Nature physics **2**(2) (2006) 110–115

[121] Naeem, S.: Species redundancy and ecosystem reliability. Conservation Biology

12(1) (1998) 39–45

[122] Petchey, O.L., McPhearson, P.T., Casey, T.M., Morin, P.J.: Environmental warming alters food-web structure and ecosystem function. Nature **402**(6757) (1999) 69–72

[123] Ledger, M., Hildrew, A.: The ecology of acidification and recovery: changes in herbivore-algal food web linkages across a stream ph gradient. Environmental Pollution **137**(1) (2005) 103–118

[124] Woodward, G.: Biodiversity, ecosystem functioning and food webs in fresh waters: assembling the jigsaw puzzle. Freshwater Biology **54**(10) (2009) 2171–2187

[125] Allesina, S., Bodini, A., Pascual, M.: Functional links and robustness in food webs. Philosophical Transactions of the Royal Society B: Biological Sciences **364**(1524) (2009) 1701–1709

[126] O'Connor, N.E., Crowe, T.P.: Biodiversity loss and ecosystem functioning: distinguishing between number and identity of species. Ecology **86**(7) (2005) 1783–1796

[127] Ma, A., Mondragón, R.: Evaluation of network robustness using a node tearing algorithm. Physica A: Statistical Mechanics and its Applications **391**(24) (2012) 6674–6681

[128] Trajanovski, S., Martín-Hernández, J., Winterbach, W., Van Mieghem, P.: Robustness envelopes of networks. Journal of Complex Networks **1**(1) (2013) 44–62

[129] Newman, M.E., Strogatz, S.H., Watts, D.J.: Random graphs with arbitrary degree distributions and their applications. Physical Review E **64**(2) (2001) 026118

[130] Holme, P., Kim, B.J., Yoon, C.N., Han, S.K.: Attack vulnerability of complex networks. Physical Review E **65**(5) (2002) 056109

[131] Dolev, D., Jamin, S., Mokryn, O.O., Shavitt, Y.: Internet resiliency to attacks and failures under bgp policy routing. Computer Networks **50**(16) (2006) 3183–3196

[132] Dunne, J.A., Brose, U., Williams, R.J., Martinez, N.D.: Modeling food-web dynamics: complexity-stability implications. Aquatic Food Webs: An Ecosystem Approach (2005) 117–129

[133] Srinivasan, U.T., Dunne, J.A., Harte, J., Martinez, N.D.: Response of complex food webs to realistic extinction sequences. Ecology **88**(3) (2007) 671–682

[134] Coll, M., Lotze, H.K., Romanuk, T.N.: Structural degradation in mediterranean sea food webs: testing ecological hypotheses using stochastic and mass-balance modelling. Ecosystems **11**(6) (2008) 939–960

[135] Kaiser-Bunbury, C.N., Muff, S., Memmott, J., Müller, C.B., Caflisch, A.: The robustness of pollination networks to the loss of species and interactions: a quantitative approach incorporating pollinator behaviour. Ecology Letters **13**(4) (2010) 442–452

[136] LaBar, T., Campbell, C., Yang, S., Albert, R., Shea, K.: Global versus local extinction in a network model of plant–pollinator communities. Theoretical Ecology **6**(4) (2013) 495–503

[137] Ramos-Jiliberto, R., Valdovinos, F.S., Moisset de Espanés, P., Flores, J.D.: Topological plasticity increases robustness of mutualistic networks. Journal of Animal Ecology **81**(4) (2012) 896–904

[138] Evans, D.M., Pocock, M.J., Memmott, J.: The robustness of a network of ecological networks to habitat loss. Ecology letters **16**(7) (2013) 844–852

[139] Maslov, S., Sneppen, K., Zaliznyak, A.: Detection of topological patterns in complex networks: correlation profile of the internet. Physica A: Statistical Mechanics and its Applications **333** (2004) 529–540

[140] Barabási, A.L., Bonabeau, E.: Scale-free networks. Scientific American **288**(5) (2003) 50–59

[141] Doyle, J.C., Alderson, D.L., Li, L., Low, S., Roughan, M., Shalunov, S., Tanaka, R., Willinger, W.: The "robust yet fragile" nature of the internet. Proceedings of the National Academy of Sciences of the United States of America **102**(41) (2005) 14497–14502

[142] W Fox, J.: Current food web models cannot explain the overall topological structure of observed food webs. Oikos **115**(1) (2006) 97–109

[143] Dormann, C.F., Fründ, J., Blüthgen, N., Gruber, B.: Indices, graphs and null models: Analyzing bipartite ecological networks. Open Ecology Journal **2** (2009) 7–24

[144] Goldstein, J., Zych, M.: What if we lose a hub? experimental testing of pollina-

tion network resilience to removal of keystone floral resources. Arthropod-Plant Interactions 1–9

[145] Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., Alon, U.: Network motifs: simple building blocks of complex networks. Science Signalling **298**(5594) (2002) 824

[146] Kruskal, J.B.: On the shortest spanning subtree of a graph and the traveling salesman problem. Proceedings of the American Mathematical society **7**(1) (1956) 48–50

[147] Atmar, W., Patterson, B.D.: The measure of order and disorder in the distribution of species in fragmented habitat. Oecologia **96**(3) (1993) 373–382

[148] Sole, R.V., Montoya, M.: Complexity and fragility in ecological networks. Proceedings of the Royal Society of London B: Biological Sciences **268**(1480) (2001) 2039–2045

[149] Benedek, Z., Jordán, F., Báldi, A.: Topological keystone species complexes in ecological interaction networks. Community Ecology **8**(1) (2007) 1–7

[150] Proulx, S.R., Promislow, D.E., Phillips, P.C.: Network thinking in ecology and evolution. Trends in Ecology & Evolution **20**(6) (2005) 345–353

[151] Moody, J., White, D.R.: Structural cohesion and embeddedness: A hierarchical concept of social groups. American Sociological Review (2003) 103–127

[152] Campbell, C., Yang, S., Shea, K., Albert, R.: Topology of plant-pollinator networks that are vulnerable to collapse from species extinction. Physical Review E **86**(2) (2012) 021924

[153] Ventresca, M., Aleman, D.: Network robustness versus multi-strategy sequential attack. Journal of Complex Networks **3**(1) (2015) 126–146

[154] Anderson, T.K., Sukhdeo, M.V.: Host centrality in food web networks determines parasite diversity. PLoS One **6**(10) (2011) e26798

[155] Wasserman, S., Faust, K.: Social network analysis: Methods and applications. Volume 8. Cambridge university press (1994)

[156] Sukhdeo, M.V., Hernandez, A.D.: Food web patterns and the parasite's perspective. Parasitism and ecosystems. Oxford University Press, Oxford (2005) 54–67

[157] Jordano, P., Bascompte, J., Olesen, J.M., et al.: The ecological consequences of complex topology and nested structure in pollination webs. Plant-pollinator interactions: from specialization to generalization (2006) 173–199

[158] Spizzirri, L.: Justification and application of eigenvector centrality. Algebra in Geography: Eigenvectors of Network (2011)

[159] Wang, Y., Di, Z., Fan, Y.: Identifying and characterizing nodes important to community structure using the spectrum of the graph. PloS one **6**(11) (2011) e27418

[160] Ghoshal, G., Barabási, A.L.: Ranking stability and super-stable nodes in complex networks. Nature communications **2** (2011) 394

[161] Milly, P.C., Dunne, K.A., Vecchia, A.V.: Global pattern of trends in streamflow and water availability in a changing climate. Nature **438**(7066) (2005) 347–350

[162] Closs, G., Lake, P.: Spatial and temporal variation in the structure of an intermittent-stream food web. Ecological Monographs (1994) 2–21

[163] Kendon, M., Marsh, T., Parry, S.: The 2010–2012 drought in england and wales. Weather **68**(4) (2013) 88–95

[164] Vörösmarty, C.J., McIntyre, P., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S.E., Sullivan, C.A., Liermann, C.R., et al.: Global threats to human water security and river biodiversity. Nature **467**(7315) (2010) 555–561

[165] Lake, P.: Ecological effects of perturbation by drought in flowing waters. Freshwater biology **48**(7) (2003) 1161–1172

[166] Petchey, O.L., Brose, U., Rall, B.C.: Predicting the effects of temperature on food web connectance. Philosophical Transactions of the Royal Society of London B: Biological Sciences **365**(1549) (2010) 2081–2091

[167] Tylianakis, J.M., Tscharntke, T., Lewis, O.T.: Habitat modification alters the structure of tropical host–parasitoid food webs. Nature **445**(7124) (2007) 202–205

[168] Brown, L.E., Edwards, F.K., Milner, A.M., Woodward, G., Ledger, M.E.: Food web complexity and allometric scaling relationships in stream mesocosms: implications for experimentation. Journal of Animal Ecology **80**(4) (2011) 884–895

[169] Bennett, L., Kittas, A., Liu, S., Papageorgiou, L.G., Tsoka, S.: Community struc-

ture detection for overlapping modules through mathematical programming in protein interaction networks. PloS one **9**(11) (2014) e112821

[170] Brede, M.: Coordinated and uncoordinated optimization of networks. Physical Review E **81**(6) (2010) 066104

[171] Derényi, I., Farkas, I., Palla, G., Vicsek, T.: Topological phase transitions of random networks. Physica A: Statistical Mechanics and its Applications **334**(3) (2004) 583–590

[172] Tilman, D., Balzer, C., Hill, J., Befort, B.L.: Global food demand and the sustainable intensification of agriculture. Proceedings of the National Academy of Sciences **108**(50) (2011) 20260–20264

[173] Flohre, A., Fischer, C., Aavik, T., Bengtsson, J., Berendse, F., Bommarco, R., Ceryngier, P., Clement, L.W., Dennis, C., Eggers, S., et al.: Agricultural intensification and biodiversity partitioning in european landscapes comparing plants, carabids, and birds. Ecological Applications **21**(5) (2011) 1772–1781

[174] Champion, G., May, M., Bennett, S., Brooks, D., Clark, S., Daniels, R., Firbank, L., Haughton, A., Hawes, C., Heard, M., et al.: Crop management and agronomic context of the farm scale evaluations of genetically modified herbicide–tolerant crops. Philosophical Transactions of the Royal Society of London B: Biological Sciences **358**(1439) (2003) 1801–1818

[175] Firbank, L., Heard, M., Woiwod, I., Hawes, C., Haughton, A., Champion, G., Scott, R., Hill, M., Dewar, A., Squire, G., et al.: An introduction to the farm-scale evaluations of genetically modified herbicide-tolerant crops. Journal of Applied Ecology **40**(1) (2003) 2–16

[176] Brooks, D.R., Clark, S.J., Perry, J.N., Bohan, D.A., Champion, G.T., Firbank, L.G., Haughton, A.J., Hawes, C., Heard, M.S., Woiwod, I.P.: Invertebrate biodiversity in maize following withdrawal of triazine herbicides. Proceedings of the Royal Society of London B: Biological Sciences **272**(1571) (2005) 1497–1502

[177] Brooks, D., Firbank, L., Bohan, D., Champion, G., Clark, S., Dewar, A., Haughton, A., Hawes, C., Heard, M., May, M., et al.: The implications of genetically modified herbicide-tolerant crops for uk farmland biodiversity: a summary of the results of

the farm scale evaluations project. In: Proceedings of the Open Seminar; Cultivation of Genetically Modified Crops and Evaluation of Ecological Effects, Tokyo November 2007. (2007) 29–52

[178] Hawes, C., Haughton, A., Bohan, D., Squire, G.: Functional approaches for assessing plant and invertebrate abundance patterns in arable systems. Basic and Applied Ecology **10**(1) (2009) 34–42

[179] Fuller, R., Gregory, R., Gibbons, D., Marchant, J., Wilson, J., Baillie, S., Carter, N.: Population declines and range contractions among lowland farmland birds in britain. Conservation Biology **9**(6) (1995) 1425–1441

[180] Chamberlain, D., Fuller, R., Bunce, R., Duckworth, J., Shrubb, M.: Changes in the abundance of farmland birds in relation to the timing of agricultural intensification in england and wales. Journal of applied ecology **37**(5) (2000) 771–788

[181] Firbank, L.: The farm scale evaluations of spring-sown genetically modified crops. introduction. Philosophical Transactions of the Royal Society B: Biological Sciences **358**(1439) (2003) 1777

[182] Haughton, A., Champion, G., Hawes, C., Heard, M., Brooks, D., Bohan, D., Clark, S., Dewar, A., Firbank, L., Osborne, J., et al.: Invertebrate responses to the management of genetically modified herbicide–tolerant and conventional spring crops. ii. within-field epigeal and aerial arthropods. Philosophical Transactions of the Royal Society of London B: Biological Sciences **358**(1439) (2003) 1863–1877

[183] Brooks, D., Bohan, D., Champion, G., Haughton, A., Hawes, C., Heard, M., Clark, S., Dewar, A., Firbank, L., Perry, J., et al.: Invertebrate responses to the management of genetically modified herbicide–tolerant and conventional spring crops. i. soil-surface-active invertebrates. Philosophical Transactions of the Royal Society of London B: Biological Sciences **358**(1439) (2003) 1847–1862

[184] Herrmann, H.J., Schneider, C.M., Moreira, A.A., Andrade Jr, J.S., Havlin, S.: Onion-like network topology enhances robustness against malicious attacks. Journal of Statistical Mechanics: Theory and Experiment **2011**(01) (2011) P01027

[185] Newman, M.E.: Mixing patterns in networks. Physical Review E **67**(2) (2003) 026126

[186] Bray, J.R., Curtis, J.T.: An ordination of the upland forest communities of southern wisconsin. Ecological monographs **27**(4) (1957) 325–349

[187] Rip, J.M., McCann, K.S., Lynn, D.H., Fawcett, S.: An experimental test of a fundamental food web motif. Proceedings of the Royal Society B: Biological Sciences (2010) rspb20092191

[188] Burgos, E., Ceva, H., Perazzo, R.P., Devoto, M., Medan, D., Zimmermann, M., Delbue, A.M.: Why nestedness in mutualistic networks? Journal of Theoretical Biology **249**(2) (2007) 307–313

[189] Sala, O.E., Chapin, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A., et al.: Global biodiversity scenarios for the year 2100. science **287**(5459) (2000) 1770–1774

[190] Raybould, A., Poppy, G.M.: Commercializing genetically modified crops under eu regulations: objectives and barriers. GM crops & food **3**(1) (2012) 9–20

[191] Iyer, S., Killingback, T., Sundaram, B., Wang, Z.: Attack robustness and centrality of complex networks. PloS one **8**(4) (2013) e59613

[192] Mulder, C., Boit, A., Mori, S., Arie Vonk, J., Dyer, S.D., Faggiano, L., Geisen, S., González, A.L., Kaspari, M., Lavorel, S., et al.: 1 distributional (in) congruence of biodiversity-ecosystem functioning. Advances in Ecological Research **46** (2012) 1

[193] Woodward, G., Speirs, D.C., Hildrew, A.G., Hal, C.: Quantification and resolution of a complex, size-structured food web. Advances in ecological research **36** (2005) 85–135

[194] Melián, C.J., Bascompte, J.: Complex networks: two ways to be robust? Ecology Letters **5**(6) (2002) 705–708

[195] Woodward, G., Papantoniou, G., Edwards, F., Lauridsen, R.B.: Trophic trickles and cascades in a complex food web: impacts of a keystone predator on stream community structure and ecosystem processes. Oikos **117**(5) (2008) 683–692

[196] Christianou, M., Ebenman, B.: Keystone species and vulnerable species in ecological communities: strong or weak interactors? Journal of Theoretical Biology **235**(1) (2005) 95–103

[197] Simberloff, D.: Flagships, umbrellas, and keystones: is single-species management passé in the landscape era? Biological conservation **83**(3) (1998) 247–257

[198] Gibson, R., Nelson, I., Hopkins, G., Hamlett, B., Memmott, J.: Pollinator webs, plant communities and the conservation of rare plants: arable weeds as a case study. Journal of Applied Ecology **43**(2) (2006) 246–257

[199] Estrada, E., Rodriguez-Velazquez, J.A.: Subgraph centrality in complex networks. Physical Review E **71**(5) (2005) 056103

[200] Amaral, L.A.N., Scala, A., Barthelemy, M., Stanley, H.E.: Classes of small-world networks. Proceedings of the national academy of sciences **97**(21) (2000) 11149–11152

[201] Everett, M.G., Borgatti, S.P.: The centrality of groups and classes. The Journal of mathematical sociology **23**(3) (1999) 181–201

[202] Dunne, J., Williams, R., Martinez, N.: Small networks but not small worlds: unique aspects of food web structure. Proc. Nat. Acad. Sci (2002)

[203] Montoya, J.M., Solé, R.V.: Small world patterns in food webs. Journal of theoretical biology **214**(3) (2002) 405–412

[204] Libralato, S., Christensen, V., Pauly, D.: A method for identifying keystone species in food web models. ecological modelling **195**(3) (2006) 153–171

[205] Klümper, W., Qaim, M.: A meta-analysis of the impacts of genetically modified crops. PLoS One **9**(11) (2014) e111629

[206] Giampietro, M.: The precautionary principle and ecological hazards of genetically modified organisms. AMBIO: A Journal of the Human Environment **31**(6) (2002) 466–470

[207] Taleb, N.N., Read, R., Douady, R., Norman, J., Bar-Yam, Y.: The precautionary principle (with application to the genetic modification of organisms). arXiv preprint arXiv:1410.5787 (2014)

[208] Opsahl, T., Agneessens, F., Skvoretz, J.: Node centrality in weighted networks: Generalizing degree and shortest paths. Social Networks **32**(3) (2010) 245–251

[209] Brose, U., Williams, R.J., Martinez, N.D.: Allometric scaling enhances stability in complex food webs. Ecology Letters **9**(11) (2006) 1228–1236

[210] Stork, N.E., Blackburn, T.M.: Abundance, body size and biomass of arthropods in tropical forest. Oikos (1993) 483–489

[211] Thompson, R.M., Brose, U., Dunne, J.A., Hall Jr, R.O., Hladyz, S., Kitching, R.L., Martinez, N.D., Rantala, H., Romanuk, T.N., Stouffer, D.B., et al.: Food webs: reconciling the structure and function of biodiversity. Trends in ecology & evolution **27**(12) (2012) 689–697

[212] Xu, G., Bennett, L., Papageorgiou, L.G., Tsoka, S.: Module detection in complex networks using integer optimisation. Algorithms for Molecular Biology **5**(1) (2010) 1

[213] De Wolf, T., Holvoet, T.: Emergence versus self-organisation: Different concepts but promising when combined. In: Engineering self-organising systems. Springer (2005) 1–15

[214] Gross, T., Blasius, B.: Adaptive coevolutionary networks: a review. Journal of The Royal Society Interface **5**(20) (2008) 259–271

[215] Willinger, W., Doyle, J.: Robustness and the internet: Design and evolution (2002)

[216] Jackson, M.O.: The stability and efficiency of economic and social networks. In: Advances in Economic Design. Springer (2003) 319–361

[217] Nagaraja, S.: Topology of covert conflict. In: Security Protocols, Springer (2007) 329–332

[218] Ochiai, H., Suzuki, R., Arita, T.: Self-organizing stability of the food web that emerges from the evolution of restrictions on speciation. Artificial Life and Robotics **16**(4) (2012) 523–525

[219] Kim, H., Anderson, R.: An experimental evaluation of robustness of networks. IEEE Systems Journal **7**(2) (2013) 179–188

[220] Tero, A., Takagi, S., Saigusa, T., Ito, K., Bebber, D.P., Fricker, M.D., Yumiki, K., Kobayashi, R., Nakagaki, T.: Rules for biologically inspired adaptive network design. Science **327**(5964) (2010) 439–442

[221] De Laender, F., Van den Brink, P., Janssen, C.: Functional redundancy and food web functioning in linuron-exposed ecosystems. Environmental Pollution **159**(10) (2011) 3009–3017

[222] Gao, F., Musial, K., Cooper, C., Tsoka, S.: Link prediction methods and their accuracy for different social networks and network metrics. Scientific Programming **2015** (2015) 1

[223] Philpott, S.M., Pardee, G.L., Gonthier, D.J.: Cryptic biodiversity effects: importance of functional redundancy revealed through addition of food web complexity. Ecology **93**(5) (2012) 992–1001

[224] Gleditsch, K.S.: Expanded trade and gdp data. Journal of Conflict Resolution **46**(5) (2002) 712–724