

Article

# Exploring the Interactions of Physical, Chemical and Biological Variables of an Urban River Using Network Analysis

Cecilia Medupin <sup>1,\*</sup>, Charles Bannister <sup>2</sup> and Jean-Marc Schwartz <sup>2</sup>

<sup>1</sup> School of Natural Sciences, The University of Manchester, Manchester M13 9PL, UK

<sup>2</sup> School of Biological Sciences, The University of Manchester, Oxford Road, Manchester M13 9PL, UK; charlesbannister@outlook.com (C.B.); Jean-Marc.Schwartz@manchester.ac.uk (J.-M.S.)

\* Correspondence: cecilia.medupin@manchester.ac.uk

Received: 5 July 2020; Accepted: 10 September 2020; Published: 15 September 2020



**Abstract:** Network analysis was used as a method to investigate the relationship between benthic macroinvertebrates in an urban river watershed and physicochemical variables. The measured physicochemical variables were the dissolved oxygen, temperature, nutrients, conductivity, pH, total organic matter, biochemical oxygen demand and river discharge. The metrics applied in the study were the degree of connections between nodes, the number of edges identified for each study location and the functional feeding groups. The river sampling took place over 14 months and sampling took place at five sites, two of which were upstream of a major wastewater treatment works and three sites were downstream of the works. A biological and environmental (BIOENV) analysis was included as part of the overall analysis to compare the variables that influenced the river ecosystem. This study shows that the relationships between benthic macroinvertebrates were stronger at the upstream locations of the watershed, while the downstream locations were controlled by the physicochemical relationships. From this analysis, the river quality and biodiversity were mainly controlled by the discharge, conductivity and availability of relevant organic matter suitable for organisms. Through the network, the degree of connections between the variables revealed the status of the urban river and provided insight into the possible management of vegetation cover across the urban watershed.

**Keywords:** network analysis; water quality; benthic macroinvertebrates; interaction; ecosystem; ecology

## 1. Introduction

Physical and chemical environmental factors, such as temperature, discharge, conductivity, nutrients and organic matter, can influence the biological components of rivers, including benthic macroinvertebrates. The metrics used to measure stream health include functional feeding groups [1–3], community diversity and the degree of connection between the physicochemical variables and invertebrates [4]. However, increasing human population growth and the associated agricultural and urban demands put large amounts of pressure on freshwater resources in terms of both the quality and quantity needed to satisfy human needs [5–12]. These pressures on rivers modify and sometimes limit the biodiversity abundance, which consequently impacts on the ecosystem services provided by these rivers [10,13,14]. The impact of point and non-point source pollution on rivers' biodiversity, richness and abundance have been widely studied [15–19]. Due to climate change, freshwater resources have become more stressed and most parts of the world [20,21], including the U.K. [11], will be challenged by changing environmental conditions. For these reasons, freshwater resources are continually influenced by multiple stressors, which means it is important to understand these stressors from an ecological perspective. To understand these multiple stressors, interactions between benthic macroinvertebrates and physicochemical variables have been studied using multivariate analysis [22–24].

In recent years, network analysis has been applied to understand complex interactions in diverse fields. For example, the method has been applied to animal social networks and disease transmission networks [25–27] in genomics as they allow for complex gene transcription factor and protein component interactions to be identified [28–30]. In psychology, network analysis has been used as an important concept and analytical approach [31], where social network analysis has been applied [32] to explain mutualisms and correlations. Thus, constituent variables can mutually influence other variables without the need to hypothesise the existence of causal latent variables [33]. In environmental management, social network analysis had been utilised to explain natural resources governance and to understand the circulation and local management of plant propagative materials. The results have shown that there are connectors involved in natural resource management, which allow for predictions to be made that might affect resource management [34].

In ecological studies, network analysis has been applied in the assessment of microbial communities using genomic indicators, e.g., [35,36]. In plant science, network analysis has been explored to progress research in above-ground and below-ground ecological interactions [37]. Network analysis has been applied in marine ecology to study the structure of energy flows between ecosystem components. Network analysis has been used to reveal major exchanges and dependencies between the major marine subsystems, and as a means to quantify the architecture of ecological complexity [4].

As a new source analysis tool in ecology, networks can link numerous components that interact with one another in a system. The combination of edges and nodes can be applied to the process of finding a solution that will be optimized such that it is interpretable via clearer physical meanings [4,31]. However, applications of this model to identify the interactions between variables in river environments have rarely been reported.

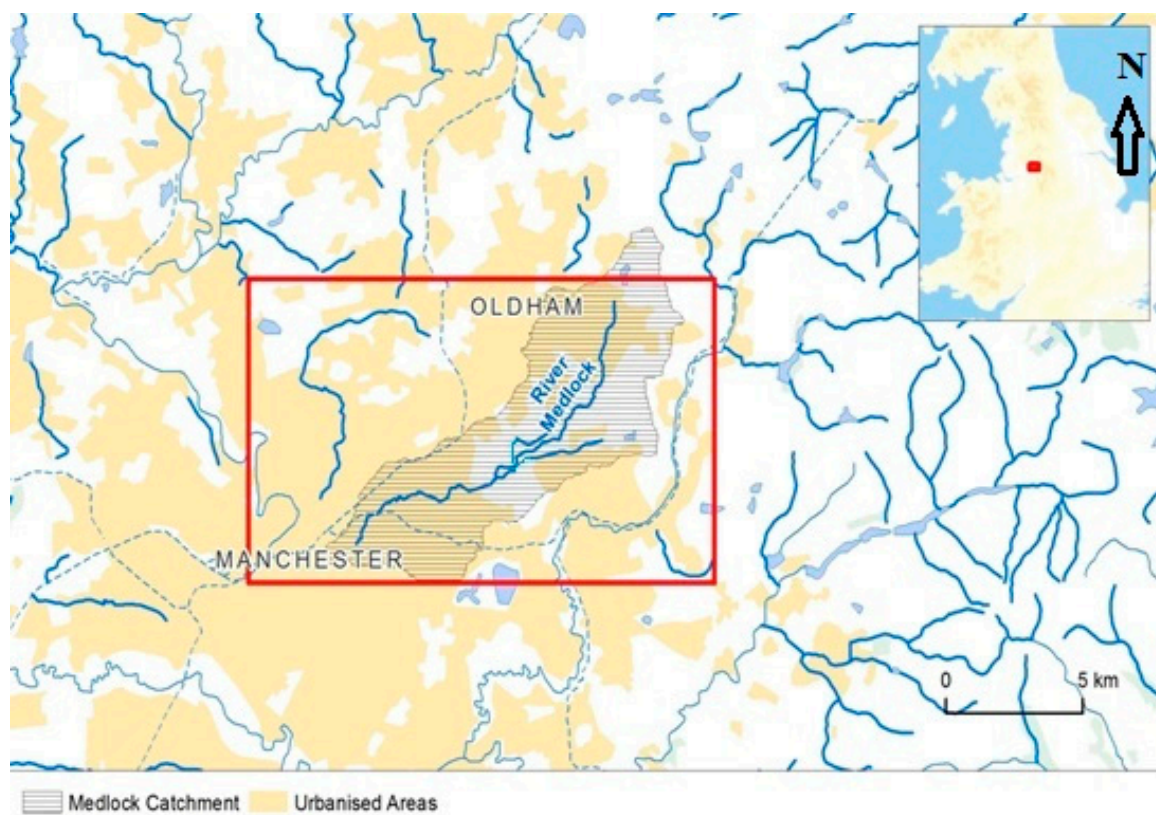
The study aimed to identify the interactions between water quality variables and benthic macroinvertebrate communities by using network analysis and to determine which environmental variable(s) influenced the urban river organisms. The results provided insight into the characteristics of the pollution sources at the Medlock River catchment and can contribute to the development of effective water quality protection strategies.

## 2. Methods

### 2.1. Study Area—The River Medlock, Greater Manchester, U.K.

The River Medlock catchment European Petroleum Survey Group (EPSG):4326 WGS 84—latitude: 53°28′26.429″, longitude: 2°13′49.691″) is one of five tributaries of the City of Manchester to the River Irwell in the northwest of England, U.K. (Figure 1). The river originates from the Pennines of the English Peak District and runs for approximately 22 km from the northeastern part of Oldham (EPSG:4326 WGS 84—latitude: 53°32′43.784″, longitude: 2°4′20.205″) to the Manchester City Centre (EPSG:4326 WGS 84—latitude: 53°28′38.112″, longitude: 2°12′56.544″). The river has a catchment size of 65.9 km<sup>2</sup>, an annual mean flow of 0.86 m<sup>3</sup>/s and a maximum altitude of 376 m. The major land cover patterns in the catchment include grassland, woodland and an urban extent (transport networks, commercial and residential areas). The river has a continuously operational wastewater treatment works (WwTW) at Failsworth EPSG:4326 WGS 84 (latitude: 53°29′41.282″, longitude: 2°9′25.615″), as well as numerous combined sewer and stormwater overflows (United Kingdom’s Environment Agency, personal communication) and surface water drains. As with other urban rivers in the Greater Manchester area, human settlements in this catchment date back as far as Roman times [38]. With increased urbanisation and the need to control water flow due to flood events, most sections of the river have been modified, with a resultant effect on the river quality [39]. The River Medlock is recorded as one of the rivers in England that have not fully complied with the “good ecological status” required by the European Union’s Water Framework Directive [40] for surface waters. A high concentration of phosphate from combined sewer overflows and wastewater treatment works is a major challenge [23]. Other challenges include channel modifications, canalisation, urban runoff, leakages

and agricultural effluent, which impact the river's diversity and abundance [39]. Further studies on the land cover for this case study can be found in [14].



**Figure 1.** River Medlock, Greater Manchester, showing catchment boundary with urban settlements. Inset: Map of the U.K. with the location of Greater Manchester.

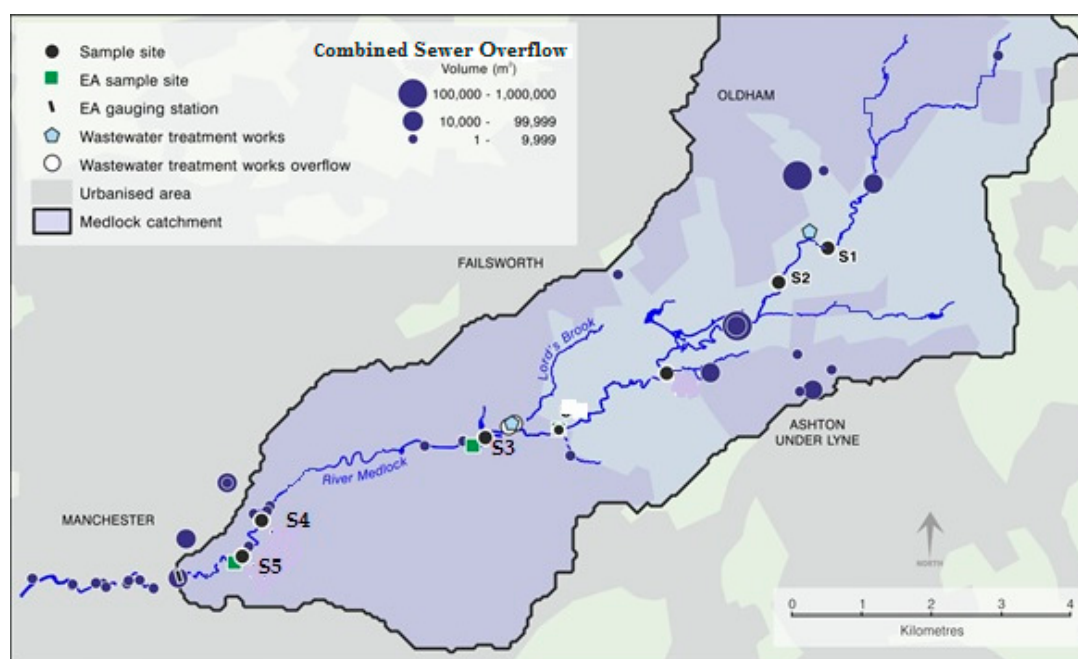
## 2.2. Samples

Water samples were obtained from March 2013 to April 2014 from 14 field campaigns carried out upstream and downstream of combined sewer overflows and the wastewater treatment works of the River Medlock (Figure 2). The sites were numbered with increasing site number based on the distance from the source following the direction of the river; S1 and S2 were above the WwTW, while S3 to S5 were below the works as the river flowed towards Manchester City Centre. The sites were also chosen by considering the temporal variation with spring (March to May), summer (June to August), autumn (September to November) and winter (from December to February). Table 1 shows the characteristics of the sites examined in the study. River substrates obtained from the river were also included in the analysis. For further information on the study locations, see [23].

Sites 1 and 2 had a lower urban cover than those located downstream of the Wastewater Treatment Works (WwTW), where sites 3–5 were considered to have increasingly urbanized catchment areas in that order. These secondary data contained 32 benthic macro-invertebrate families, which constituted 13,664 invertebrate sample counts and physicochemical variables made up of 770 samples (14 months  $\times$  5 sites  $\times$  11 variables) obtained from S1 to S5.

Data was measured for physicochemical variables, including pH; dissolved oxygen (percentage saturation); conductivity ( $\mu\text{Scm}^{-1}$ ); river discharge ( $\text{m}^3 \text{s}^{-1}$ ); temperature ( $^{\circ}\text{C}$ ); concentrations (mg/L) of nitrate-N (nitrate), phosphate-P (phosphate), total organic matter (TOM), ammonia-N (ammonia), biological oxygen demand (BOD) and suspended solids. All variables were analysed according to the standard analytical requirements [41]. The benthic macroinvertebrate samples were obtained through the standard three-minute kick sampling techniques and identified at the taxonomical family level [42].

Other physical variables in Table 1, including the catchment area, altitude and slope, were determined using information obtained from the ArcGIS Geoprocessing toolbox [43].



**Figure 2.** Map of each of the five sites along the River Medlock for which invertebrate sample counts and physicochemical readings were provided in a secondary dataset. The river flowed in a south-westerly direction towards Manchester City Centre. Sites 1 and 2 had a more rural catchment area, situated further from the city centre with increasing degrees of urbanisation when approaching the city centre. Sites 3, 4 and 5 were situated downstream of the Failsworth Wastewater Treatment Works (WwTW), which was a major point source of pollution. Created using ArcGIS 10.4.1.

**Table 1.** Site characteristics of the river Medlock.

Variables	S1	S2	S3	S4	S5
Distance (km) from source	6.6	8.5	13	16.1	17.4
Catchment area (km <sup>2</sup> )	13.65	20.55	43.98	50.16	50.84
Average slope (%)	10.72	9.45	7.56	7.24	7.21
Altitude (m)	140	107	78.9	31.1	33.6
Urbanisation (%)	3.7	10.5	10.5	12.3	12.3

### 2.3. Data Analysis

The physicochemical variables and benthic macroinvertebrates for each taxonomical family were normalised to allow for a comparison of the counts between sample dates and sites. The abundance values were normalised by dividing the number of organisms counted for one family by the total number of all organisms found at one site for one date.

To produce an interpretable and ecologically informative network graph, a minimum correlation coefficient ( $\rho$ ) threshold required to remove lower-scoring associations was set and established by testing a range of correlation coefficients, namely,  $\rho < 0.4$ ,  $\rho > 0.6$  and  $\rho = 0.9$ . At  $\rho < 0.4$ , the threshold value was found to be too low and showed many edges in the merged network; at  $\rho = 0.9$ , it was observed that the threshold value was too high and returned a merged network with only 10 edges, and thus, the graph was uninterpretable. A final threshold value of  $\rho > 0.6$  was chosen to obtain the edges recorded in the final merged network between and within each site. The mean percentage composition and standard deviation of each species for each sampling date and site were also calculated. All initial data

processing and analysis were performed using RStudio Team (2015) (RStudio: Integrated Development for R. RStudio, Inc., Boston, MA, USA (Computer Software v0.98.1074) <http://www.rstudio.com/>).

### 2.3.1. Creation of the Network Using Ecolo\_Works Pipeline

The Ecolo\_works pipeline is composed of two novel command-line Python programs, namely, NetworkCreate and NetworkMerger, which were designed for network creation and visualisation. NetworkCreate produces a network file based on a Spearman's rank correlation coefficient matrix between and within variables, and the NetworkMerger combines multiple network files that contain information related to the number of repeated edges between networks. For access to the package programs and full usage instructions, please visit [https://github.com/charlesjban/ecolo\\_works](https://github.com/charlesjban/ecolo_works). These files were loaded into the network visualisation software Cytoscape [44,45] to produce a network graph.

### 2.3.2. NetworkCreate and NetworkMerger

The program NetworkCreate.py takes as input both the correlation coefficient matrix (-cfile argument) and the corresponding  $p$ -value matrix (-pfile argument). The program outputs a network file in .csv format, which is a tab-separated text file with each line representing a network edge. While the nodes (physicochemical and biological variables) were represented using rectangles, the edges were represented using solid lines. Users have the option to define a minimum threshold value (cvalue argument, default = 0.5) for both the correlation coefficient ( $\rho$ ) and  $p$ -value ( $-p$  value argument, default = 0.05). An additional feature of the program allows users to assign a "type" label for each variable in an additional text file (-varfile argument). NetworkCreate uses this file to label the "interaction type" for each edge and creates an output "node type" file containing a tab-separated table of each node and the respective variable type. Each interaction must surpass thresholds for both the correlation coefficient and  $p$ -values to be included in the output network and shows whether the correlation is positive or negative. NetworkCreate was used to build a network for each study site based on the coefficient matrix and the corresponding  $p$ -value matrix for each study site.

NetworkMerger.py was designed to take up to nine "network" files (created using NetworkCreate.py) that are obtained as aggregates for the sampling duration, which are then merged into a single network. This network shows the interaction type and the direction of the correlation by indicating where conflicting positive or negative correlations are found and shows where they are different in the networks. The nodes were highlighted in blue for the physicochemical variables and green for biological variables. The edges were coloured grey and red for positive and negative correlations, respectively. In the merged network, the edge line thickness increased for variables with repeated numbers of interactions at different sites. The resulting network files of both NetworkCreate and NetworkMerger (Ecolo\_works) were designed to be loaded into network visualisation software, such as Cytoscape [44] and used for visualisation, where variables, such as nodes, edges, shapes and colours can be changed. The loading of the "node type" file assists in the conditional visual formatting based on the node variable type. All networks produced by NetworkCreate for each site were then merged using NetworkMerger to create a single network.

### 2.3.3. Biological and Environmental (BIOENV) Analysis

The biological and environmental (BIOENV) analysis was used to determine which variable(s) affected the benthic invertebrate communities' abundance and distribution. The BIOENV analysis was based on a weighted Spearman rank correlation coefficient ( $\rho$ ) between the physicochemical variables and benthic invertebrate communities. Both the environmental and biological variables were transformed and normalised to allow for a comparison at the same scale. The rank with the largest  $\rho$  was taken to identify the best match with the benthic macroinvertebrate communities. The BIOENV analysis was performed using Primer v6 [46] and the outcome was used to provide a comparison with

the network analysis. The river quality was classified under the European Union Water Framework Directive (EU WFD standards [47]).

### 3. Results

#### 3.1. Benthic Macroinvertebrates and Physicochemical Variables

A total of 32 benthic macroinvertebrate families were identified at the River Medlock. The dominant families included Baetidae, Chironomidae and Tubificidae. Other taxonomical families identified in limited abundance were Heptageniidae, Gammaridae, Simuliidae, Lumbriculidae, Tipulidae and Paediciidae. Table 2 shows that the mean number of families found in the river was the highest at site 1 ( $7.769 \pm 3.270$ ) and the lowest was at site 5 ( $4.50 \pm 3.28$ ). The mean organism count per site sampling peaked at S3. Five major functional feeding groups (FFGs) were identified for the taxa [1,48,49], including collector-filterers (CFs), predators (Prs), shredders (Shs), collector-gatherers (CGs) and scrapers (Scs). The river was dominated by CFs, with the highest numbers recorded at S1 and S3; scrapers were found only at S1 and S2.

**Table 2.** Benthic macroinvertebrate assemblages identified at S1 to S5, including the mean  $\pm$  SD for each family identified per sample, total organisms identified per sample and the dominant group for each study site. The families were also classified based on their functional feeding groups (FFGs): collector-filterer (CF), collector-gatherer (CG), predator (Pr), shredder (Sh) and scraper (Sc).

Sites	Mean Number of Families Present per Sampling	Mean Organism Count per Sampling	Most Dominant Families (Mean % of Total Organism Abundance)
S1	$7.769 \pm 3.270$	$63.62 \pm 46.31$	Baetidae ( $22.25 \pm 19.29$ ) (CG) Heptageniidae ( $16.88 \pm 17.52$ ) (Sc) Chironomidae ( $12.09 \pm 20.49$ ) (CF) Tubificidae ( $8.49 \pm 17.84$ ) (CF) Simuliidae ( $7.50 \pm 16.46$ ) (CF) Lumbriculidae ( $7.31 \pm 17.31$ ) (CF)
S2	$6.29 \pm 2.61$	$65.71 \pm 57.11$	Baetidae ( $30.30 \pm 31.45$ ) (CG) Chironomidae ( $19.34 \pm 23.48$ ) (CF) Tubificidae ( $8.33 \pm 21.39$ ) (CF) Heptageniidae ( $7.73 \pm 08.07$ ) (Sc) Tipulidae ( $6.25 \pm 11.83$ ) (Sh) Paediciidae ( $2.72 \pm 4.70$ ) (Pr)
S3	$6.69 \pm 2.25$	$74.38 \pm 62.97$	Baetidae ( $34.74 \pm 28.22$ ) (CG) Tubificidae ( $19.02 \pm 24.87$ ) (CF) Chironomidae ( $11.47 \pm 11.50$ ) (CF) Lumbriculidae ( $9.31 \pm 15.87$ ) (CF) Lumbriculidae ( $5.31 \pm 6.08$ ) (CF) Gammaridae ( $4.09 \pm 8.10$ ) (Pr)
S4	$6.29 \pm 2.49$	$44.38 \pm 33.14$	Baetidae ( $31.22 \pm 24.03$ ) (CG) Gammaridae ( $25.44 \pm 20.89$ ) (Pr) Chironomidae ( $14.99 \pm 25.20$ ) (CF) Lumbriculidae ( $5.38 \pm 12.52$ ) (CF) Tipulidae ( $4.52 \pm 8.06$ ) (Sh) Lumbriculidae ( $3.27 \pm 6.08$ ) (CF)
S5	$4.50 \pm 3.28$	$40.86 \pm 45.70$	Chironomidae ( $25.81 \pm 29.57$ ) (CF) Gammaridae ( $20.48 \pm 26.32$ ) (Pr) Baetidae ( $17.34 \pm 26.15$ ) (CG) Tubificidae ( $15.96 \pm 34.15$ ) (CF) Heptageniidae ( $9.03 \pm 27.47$ ) (Sc) Lumbriculidae ( $4.42 \pm 10.08$ ) (CF)

Excluding temperature ( $^{\circ}\text{C}$ ), dissolved oxygen (% saturation) and pH, the physicochemical variables of conductivity ( $\mu\text{Scm}^{-1}$ ), suspended solids ( $\text{mg L}^{-1}$ ), discharge ( $\text{m}^3 \text{s}^{-1}$ ),  $\text{BOD}_5$  ( $\text{mg L}^{-1}$ ), ammonia ( $\text{mg L}^{-1}$ ), nitrate ( $\text{mg L}^{-1}$ ), phosphate ( $\text{mg L}^{-1}$ ) and total organic matter (TOM) had higher concentrations from S3 to S5. The variables pH, temperature, dissolved oxygen,  $\text{BOD}_5$  and ammonia ( $\text{mg L}^{-1}$ ) complied with the EU's Water Framework Standard for surface waters, while phosphate ( $>0.1 \text{ mg/L}$ ) did not [50].

### 3.2. Physicochemical Variables and Benthic Macroinvertebrates

The number of interactions between the biological—biological, biological—physicochemical, physicochemical—physicochemical variables are presented in Table 3, including the number of nodes and edges. Figure 3 shows the individual networks per site and Figure 4 shows the merged networks. At S1, the interactions between the physicochemical variables and the benthic macroinvertebrates were both negative and positive (Figure 3), but at S2 to S5, these were mainly positive. Given these relationships, S1 had the highest number of interactions (34 edges) between the networks, while S5 had the lowest number of interactions (22 edges). The average number of edges for the interactions between the benthic invertebrates—benthic invertebrates was 12 edges, the benthic invertebrates—physicochemical interactions had 7.2 edges and the physicochemical—physicochemical interactions had 6.2 edges. For the merged networks, a total of 43 nodes and 106 edges were recorded in the merged network (Figure 4). Overall, the network consisted of a highly connected component of benthic invertebrates—benthic invertebrates' interactions with an average of 12 edges. These interactions showed that the organisms tended to interact with each other in the food chain based on the availability (concentrations) of physicochemical variables within the river. These interactions could be derived from taxa sharing similar ecological niches.

**Table 3.** Network node and edge counts for each site's network based on Spearman's correlation coefficient of  $\rho > 0.6$  and where  $p < 0.05$ . The counts include the total number of nodes, the total number of edges between each variable and the total number of edges for each site. The positive (+ve) and negative (−ve) correlations are given, as well as the types of edges found at the study sites (B–B: biological—biological, B–PC: biological—physicochemical and P–P: physicochemical—physicochemical). BOD: biological oxygen demand, TOM: total organic matter.

Sites	Total Nodes	B–B Edges	B–PC Edges	PC–PC Edges	Total Edges	Correlations
S1	29	15	14	5	34	Temperature = −ve dissolved oxygen, ammonia, Lumbricidae and Simuliidae Ammonia-N = −ve Perlodidae Discharge = −ve Gammaridae pH = −ve Pediciidae and conductivity Conductivity = −ve BOD BOD = −ve Ephemereallidae
S2	31	18	7	3	29	Temperature = +ve phosphate, Heptageniidae and Dytiscidae pH = −ve Pediciidae and conductivity Conductivity = −ve BOD BOD = −ve Ephemereallidae Discharge = +ve TOM & Ammonia-N pH = +ve Tubificidae and Hydropsychidae BOD = +ve Ephemereallidae Dissolved oxygen = +ve Tipulidae

Table 3. Cont.

Sites	Total Nodes	B–B Edges	B–PC Edges	PC–PC Edges	Total Edges	Correlations
S3	22	10	3	9	22	Temperature = +ve dissolved oxygen, phosphate and nitrate pH = +ve ammonia Dissolved oxygen = +ve phosphate and nitrate Ammonia = +ve phosphate, suspended solids and pH Conductivity = +ve Gammaridae and Hydropsychidae
S4	24	6	8	7	21	Temperature = +ve phosphate, nitrate and dissolved oxygen Ammonia = +ve pH and discharge Conductivity = +ve Gammaridae Phosphate = +ve dissolved oxygen, Gammaridae and Lumbriculidae BOD = +ve Simuliidae Discharge = +ve Tipulidae, Erpobdellidae and Polycentropodidae
S5	23	11	4	7	22	Temperature = +ve phosphate, nitrate and dissolved oxygen Ammonia = +ve conductivity BOD = +ve pH and Lumbriculidae Discharge = +ve TOM and Baetidae Suspended solids = +ve Baetidae Dissolved oxygen = +ve Tubificidae
Average	25.8	12	7.2	6.2	25.6	

Table 4 shows the physicochemical variables that correlated with the benthic macroinvertebrate families based on the number of correlations in the following descending order: discharges (5 families), dissolved oxygen (5), BOD (4), temperature (4), pH (3), conductivity (2), phosphate (2), ammonia (1) and suspended solids (1). After analysing the functional composition of the families, it was found that the collector-filterers were the most abundant, with a 40% contribution dominated by Oligochaetes, Simuliidae, Hydropsychidae and Polycentropodidae. The predators ranked second with 27% represented by the following families: Dytiscidae, Gammaridae and Erpobdellidae. Shredders and collector-gatherers both ranked third with a 13% contribution each, while scrapers contributed the least to the group with an abundance of 7%. Small variations were found between sites, as shown in Table 2. Predators were found at all sites except site 1; scrapers were found at sites 1, 2 and 5; shredders were found at sites 2 and 4.

**Table 4.** Nodes (benthic macroinvertebrates) and the correlated edges (physicochemical variables) for the studied river. The numbers in brackets signify the connections between the nodes and the individual edges. The families were also classified based on their functional feeding groups (FFGs): collector-filterer (CF), collector/gatherer (CG), predator (Pr), shredder (Sh) and scraper (Sc).

Discharge (5)	Dissolved Oxygen (5)	BOD (4)	Temperature (4)	pH (3)	Conductivity (2)	Ammonia (1)
Baetidae (CG)	Lumbriculidae (CF)	Ephemerellidae (CG)	Lumbriculidae (CF)	Pediciidae (Pr)	Gammaridae (Pr)	Perlotidae (Sh)
Gammaridae (Pr)	Simuliidae (CF)	Tubificidae (CF)	Simuliidae (CF)	Tubificidae (CF)	Hydropsychidae (CF)	
Tipulidae (Sh)	Dytiscidae (Pr)	Simuliidae (CF)	Dytiscidae (Pr)	Hydropsychidae (CF)	<b>Phosphate (2)</b>	<b>Suspended Solids (1)</b>
Erpobdellidae (Pr)	Tipulidae (Sh)	Lumbriculidae (CF)	Heptageniidae (Sc)		Gammaridae (Pr)	Baetidae (CG)
Polycentropodidae (CF)	Tubificidae (CF)				Lumbriculidae (CF)	



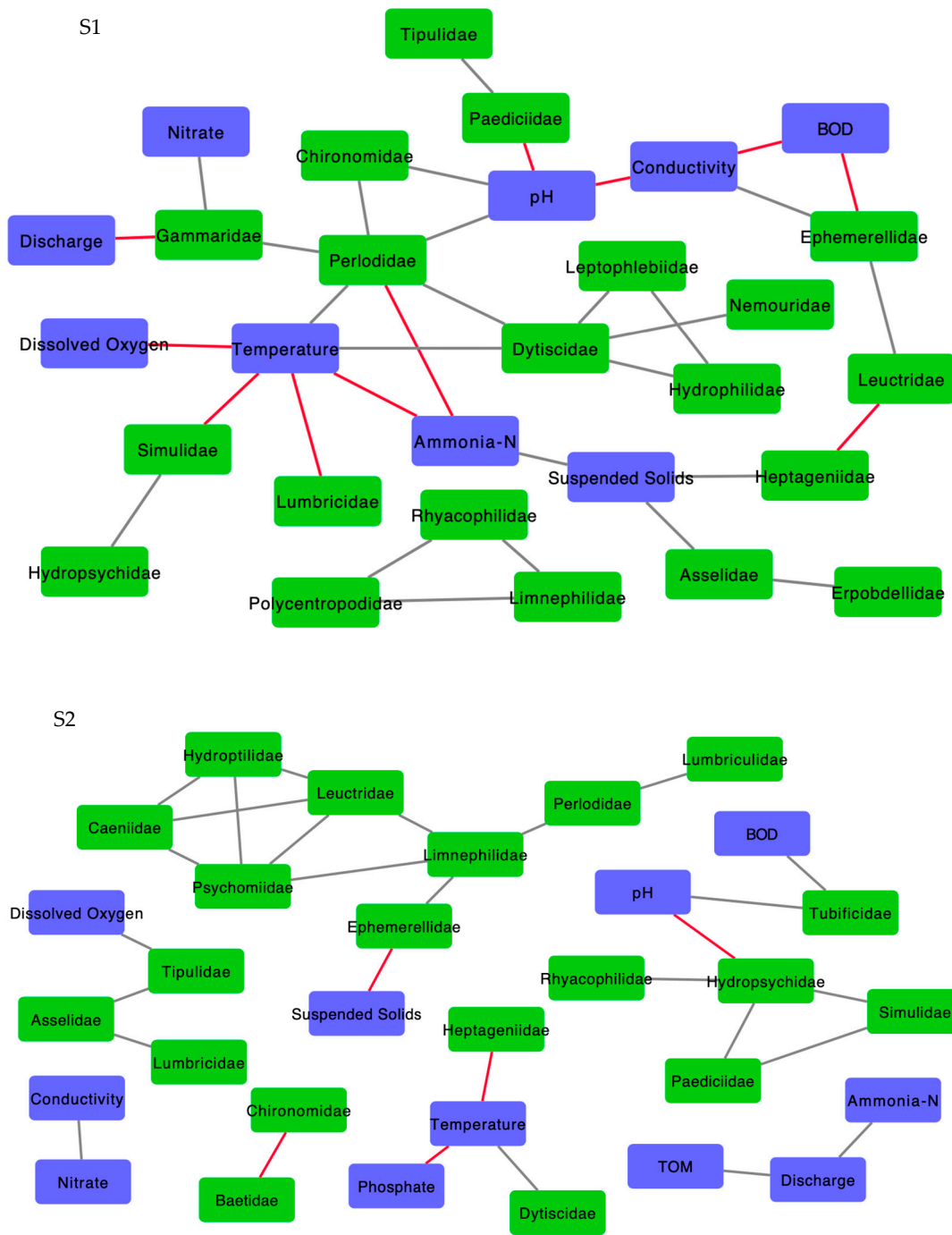


Figure 3. Cont.

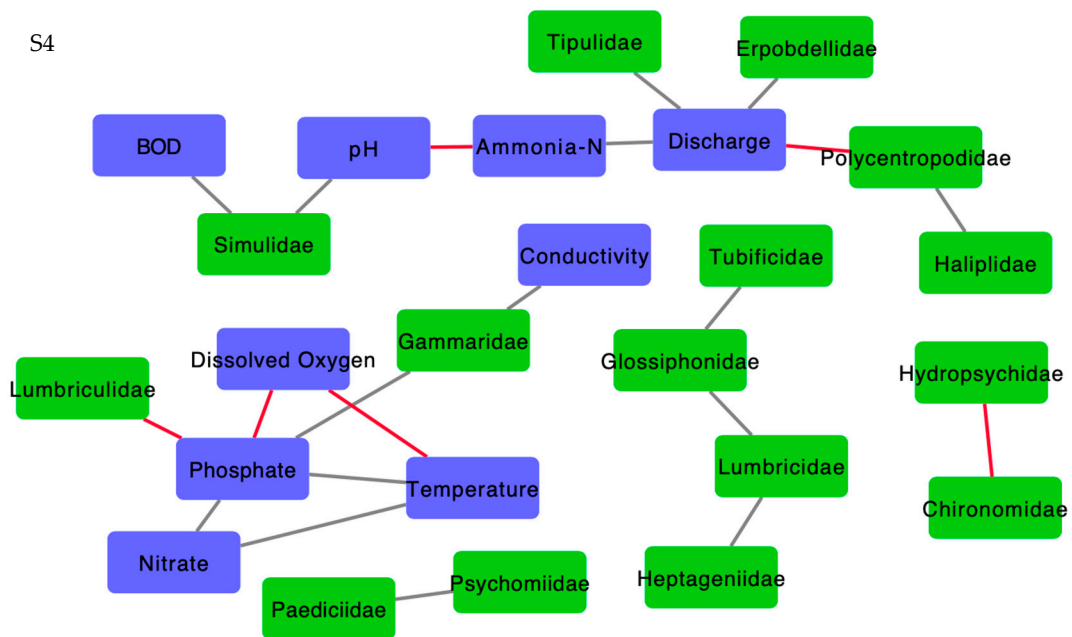
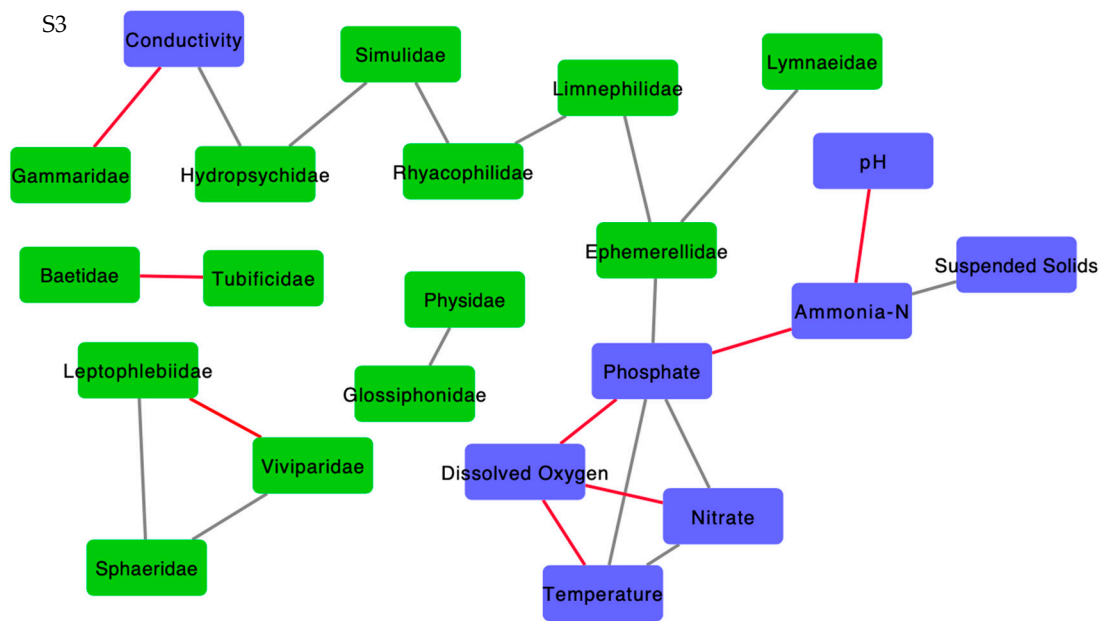
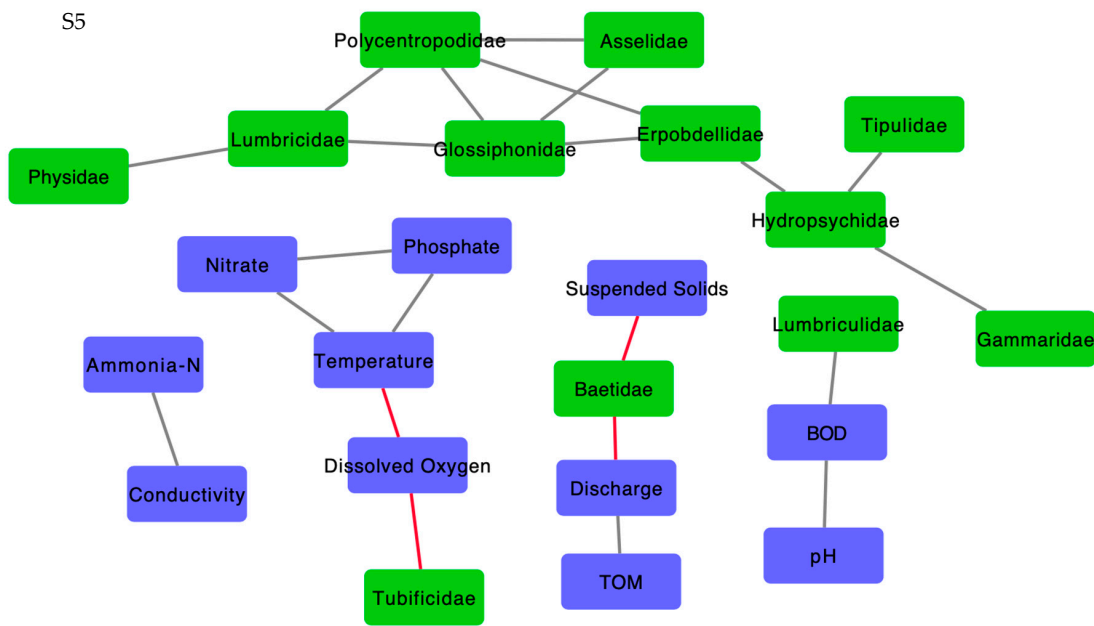
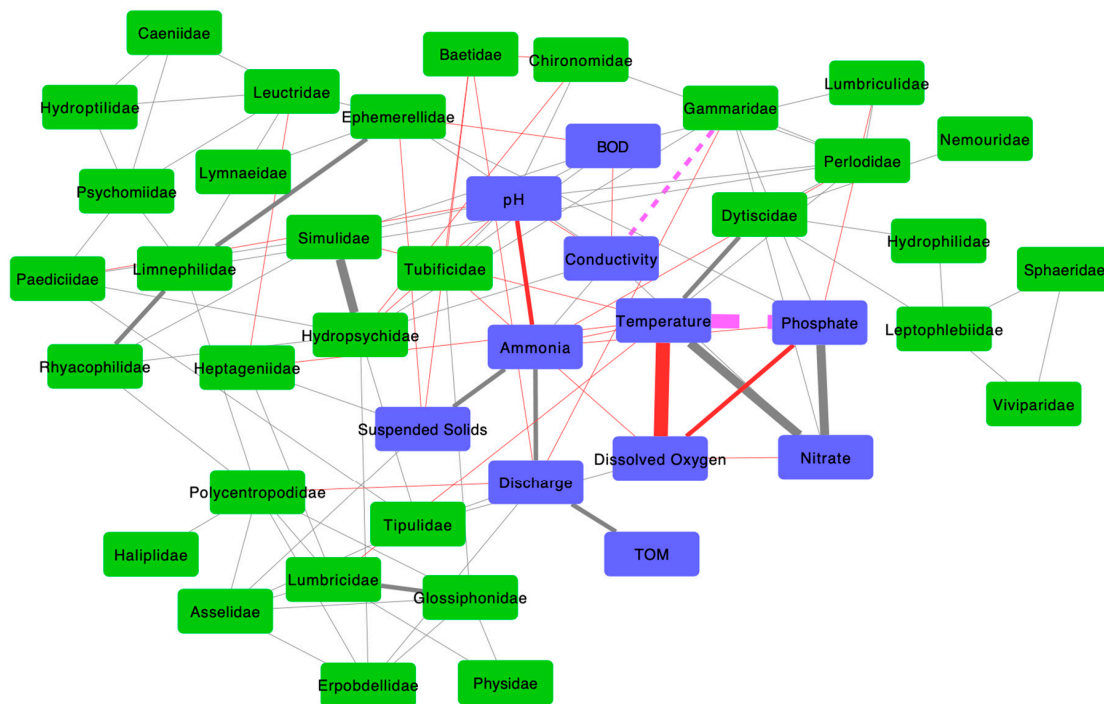


Figure 3. Cont.



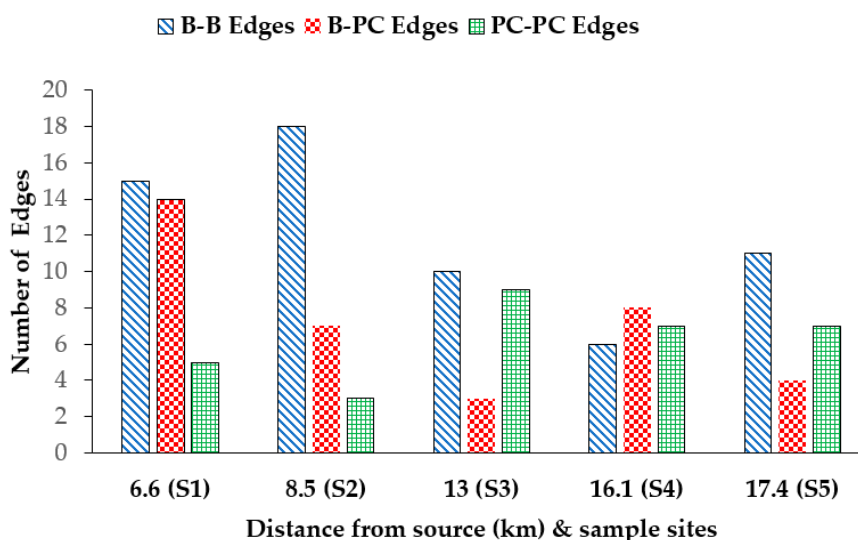
**Figure 3.** Visualised network of the Spearman’s rank correlation coefficients for sites 1 to 5 aggregated over the sampling period. Each node represents a variable at the site, where blue nodes represent physicochemical variables and green nodes represent macroinvertebrate counts. Each edge represents a significant ( $p < 0.05$ ) correlation coefficient value between the variables, where  $\rho > 0.6$ . The networks were created with the novel NetworkCreate.py program using a correlation matrix for each site, which was created using the R rcorr function [51].



**Figure 4.** Visualised network of the Spearman’s rank correlation coefficients for sites 1 to 5 (a merged network). Each node represents a variable at the site, where blue nodes represent physicochemical variables and green nodes represent macroinvertebrate counts. The grey and red edges represent positive and negative interactions, respectively. Each edge represents a significant ( $p < 0.05$ ) correlation coefficient value between the variables, where  $\rho > 0.6$ . Edges that were found at more than one site have thicker lines.

The interactions displayed in Figure 4 show thick, stronger edges for biological–biological variables, including Simuliidae and Hydropsychidae, and between Rhyacophilidae, Limnephillidae and Ephemerellidae. Among these organisms, Hydropsychidae and Simuliidae were the only organisms with repeated interactions at more than two sites, correlating positively with  $\rho = 0.69$ ,  $\rho = 0.61$  and  $\rho = 0.63$  at S1 to S3, respectively. Among the taxa groups, the insects had the strongest interactions.

The biological–biological relationship between the benthic macroinvertebrates was the highest at site 2. Site 3 had biological–physicochemical as its lowest interaction and physicochemical–physicochemical as its highest interaction (Figure 5).



**Figure 5.** The number of edges observed between the variables per site for sites 1 to 5. With increasing distance from the river’s source, the number of edges decreased at the sites. B–B: biological–biological, B–PC: biological–physicochemical and PC–PC: physicochemical–physicochemical.

### 3.3. Physicochemical–Physicochemical Relationships

Table 3 shows that five variables had a major influence on the physicochemical conditions of the study area in the following descending order: ammonia concentration, temperature, dissolved oxygen, BOD and discharge. The ammonia concentration correlated with six variables (suspended solids, pH, discharge, phosphate concentration, conductivity and temperature), the temperature correlated with four variables (i.e., dissolved oxygen and the phosphate, nitrate and ammonia concentrations), the dissolved oxygen correlated with three variables (phosphate and nitrate concentrations and temperature), the BOD correlated with two variables (pH and conductivity) and the discharge correlated with two variables (TOM and ammonia concentration).

Figures 4 and 5 show that the discharge correlated positively with ammonia concentration at sites 2 and 4 ( $\rho = 0.78$  and  $\rho = 0.63$ , respectively), and with the total organic matter concentration for the same sites ( $\rho = 0.69$  and  $\rho = 0.69$ , respectively). The ammonia concentration was also positively associated with the suspended solids concentration at sites 1 and 3 ( $\rho = 0.71$  and  $\rho = 0.63$ , respectively), and negatively associated with pH at sites 3 and 4 ( $\rho = -0.64$  and  $\rho = -0.65$ , respectively). A negative correlation between the temperature and dissolved oxygen levels was recorded at all sites, except site 2 (S1, S3 to S5:  $\rho = -0.70$ ,  $\rho = -0.80$ ,  $\rho = -0.89$  and  $\rho = -0.65$ , respectively). The temperature and phosphate concentration were negatively associated at site S2 ( $\rho = -0.63$ ). At sites 3 to 5, positive associations were seen between the temperature and phosphate concentration ( $\rho = 0.77$ ,  $\rho = 0.74$  and  $\rho = 0.72$ , respectively), between the temperature and nitrate concentration ( $\rho = 0.83$ ,  $\rho = 0.64$  and  $\rho = 0.77$ , respectively) and the temperature and phosphate concentrations ( $\rho = 0.84$ ,  $\rho = 0.75$  and  $\rho = 0.90$ , respectively). The individual networks at each site enabled the interactions to be viewed

clearly, while the merged networks (Figure 4) showed repeated edges, especially for variables that occurred frequently.

### 3.4. BIOENV Analysis

By applying the river characteristics, including the catchment area, altitude and slope, together with all the physicochemical variables and benthic macroinvertebrates (Section 3.1), Table 5 shows that the conductivity, discharge and catchment area were the variables that significantly correlated with the benthic macroinvertebrate communities, with a correlation coefficient of  $\rho = 0.274$ . However, with a correlation coefficient of  $\rho = 0.272$  (0.002 less than 0.274), the BIOENV results indicated that conductivity and catchment area were the strongest abiotic variables.

**Table 5.** Weighted Spearman's rank correlation between the biotic and abiotic variables using the BIOENV procedure.

Number of Variables	Weighted Spearman's Rank ( $\rho$ )	Variables
5	0.274	Conductivity, discharge, catchment area, altitude and slope
4	0.273	Conductivity, discharge, catchment area and altitude
5	0.272	Conductivity, phosphate, discharge, catchment area and altitude
2	0.272	Conductivity and catchment area
5	0.271	Conductivity, nitrate, discharge, catchment area and altitude

## 4. Discussion and Conclusions

The discharge, dissolved oxygen, conductivity and catchment areas were found to be the major variables impacting the river (Sections 3.2–3.4). The discharge had a positive effect on the river ecosystem [52,53] due to the transport of drift, organic matter and food sources for benthic invertebrates and the dilution of contaminants. Although the differences between the upstream (higher altitude) and downstream (lower altitude) locations in the study areas were small (31–140 m), changes in the concentrations between the upper and lower sites were noted. For example, higher levels of conductivity, nutrients and BOD were recorded at the downstream locations (S3–S5), which suggest the impact of increased urban land cover and catchment contributions [47] from the wastewater treatment works' effluent and the intermittent discharge from combined sewer overflows.

The upstream section of the river had higher counts of benthic invertebrates when compared to the lower sections of the river. The distribution of the functional feeding groups (FFGs) (Tables 2 and 4) showed the dominance of the collector-filterers (40%), which suggested the role of discharge in transporting drift that served as food to the groups [1], indicating the absence of non-forest covers and increased urban cover [14], especially at the more urban sections of the river. Windsor et al. [54] noted that contaminated rivers found in the highly urban sites of South Wales, U.K., were characterised by their reduced taxonomic and functional diversity and simplified food web structure. The Song Stream watershed in Korea [55] showed that streams close to non-vegetative sources had lower taxa richness and abundance, and therefore, were dominated by tolerant colonisers, such as Tubificidae and Hydropsychidae, which dominated the stream due to catchment contributions and sand deposition. The role of catchment contributions in the abundance of collector-filterers was also reported by [56] in a study carried out in a tropical lagoon in West Africa. While the disappearance of scrapers could be linked to the high level of organic matter and other contaminants, some studies [57,58] have shown that scrapers are intolerant of low water quality and high nutrient concentrations. In this study, the scrapers were found at S1 and S2, where the physicochemical concentrations were relatively low

but still higher at the downstream sites. The near-equal distribution of collector-gatherers, predators and shredders, along with the abundance of collector-filterers across the watershed, showed that the river was impacted by diverse environmental variables. These were impacted by the discharge, nutrient input, suspended solids, temperature, contribution of contamination from the surrounding watershed, organic matter processing [59] and the role of intermittent discharge in the transport of drift.

Whilst most studies tend to focus on interspecies interactions [60], this study provided a framework using interfamily and physicochemical relationships to visualise the interactions between and within groups of biological variables and physicochemical variables in the same co-occurrence network in a river ecosystem. The results also successfully identified strong associations between physicochemical variables at three of the five study sites tested in the study. One metric used in the study was the degree of connection in terms of the number of nodes [4], which were linked by edges (Table 4). Through the metric, the research revealed that variables that controlled the river quality and invertebrate abundance were the discharge, dissolved oxygen, BOD and temperature. These physicochemical variables were further highlighted by the number of edges that connected the nodes for each sample site (Figure 5), showing the impact of the physicochemical—physicochemical contribution to the river, especially at the downstream sections. By comparing the variables identified through network and BIOENV analyses, this study found that the variable that influenced the river concentrations at the study sites was the catchment area.

When used with other metrics, such as diversity and functional feeding groups, network analysis provides rich insight into river ecosystems, and in particular, the urban rivers. Network analysis can link many types of data across levels; therefore, it can relate individual taxa at the community or ecosystem levels. These analyses allow for the top ecosystem interactions to be highlighted by removing the low-scoring correlations from the network to provide insight into the sensitivity of the organisms to the river quality and other stressors and to identify subtle interactions. For this reason, network analysis is a powerful tool for analysing complex data relationships, and if considered in terms of ecological metrics [4], it could open up the use of the method in ecological analysis.

Furthermore, through the network visualisation, the relationships between variables are clearer and appealing [61]. The network outputs could provide opportunities for sharing scientific information with diverse groups of people, e.g., during public engagement, for evidence-based advocacy required by decision-makers and with other researchers. Therefore, network analysis has the potential to contribute to a wide range of ecological studies and provides an efficient scientific method as a catchment-based approach to river management.

**Author Contributions:** Conceptualization, C.M. and J.-M.S.; data curation, C.M. and C.B.; investigation, C.M., C.B. and J.-M.S.; methodology, C.B. and J.-M.S.; software, C.B. and J.-M.S.; supervision, C.M. and J.-M.S.; validation, J.-M.S.; writing—original draft, C.M.; writing—review and editing, C.M. and J.-M.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ramírez, A.; Gutiérrez-Fonseca, P.E. Functional feeding groups of aquatic insect families in Latin America: A critical analysis and review of existing literature. *Rev. Biol. Trop.* **2014**, *62*, 155–167. [[CrossRef](#)] [[PubMed](#)]
2. Kenney, M.; Sutton-Grier, A.; Smith, R.; Gresens, S. Benthic macroinvertebrates as indicators of water quality: The intersection of science and policy. *Terr. Arthropod Rev.* **2010**, *2*, 99–128.
3. Morley, S.A.; Karr, J.R. Assessing and restoring the health of urban streams in the Puget Sound Basin. *Conserv. Biol.* **2002**, *16*, 1498–1509. [[CrossRef](#)]
4. Lau, M.K.; Borrett, S.R.; Baiser, B.; Gotelli, N.J.; Ellison, A.M. Ecological network metrics: Opportunities for synthesis. *Ecosphere* **2017**, *8*, e01900. [[CrossRef](#)]
5. Chin, A. Urban transformation of river landscapes in a global context. *Geomorphology* **2006**, *79*, 460–487. [[CrossRef](#)]

6. Dudgeon, D. Responses of benthic macroinvertebrate communities to altitude and geology in tributaries of the Sepik River (Papua New Guinea): The influence of taxonomic resolution on the detection of environmental gradients. *Freshw. Biol.* **2012**, *57*, 1794–1812. [[CrossRef](#)]
7. Francis, R.A. Positioning urban rivers within urban ecology. *Urban. Ecosyst.* **2012**, *15*, 285–291. [[CrossRef](#)]
8. Gregory, K.J. The human role in changing river channels. *Geomorphology* **2006**, *79*, 172–191. [[CrossRef](#)]
9. McGrane, S.J. Impacts of urbanisation on hydrological and water quality dynamics, and urban water management: A review. *Hydrol. Sci. J.* **2016**, *61*, 2295–2311. [[CrossRef](#)]
10. Walsh, C.J.; Roy, A.H.; Feminella, J.W.; Cottingham, P.D.; Groffman, P.M.; Morgan, R.P. The urban stream syndrome: Current knowledge and the search for a cure. *J. N. Am. Benthol. Soc.* **2005**, *24*, 706. [[CrossRef](#)]
11. Hunter, J. *Natural Assers North: Water in the Northern Powerhouse*; Institute for Public Policy Research: Manchester, UK, 2019.
12. Hatt, B.E.; Fletcher, T.D.; Walsh, C.J.; Taylor, S.L. The influence of urban density and drainage infrastructure on the concentrations and loads of pollutants in small streams. *Environ. Manag.* **2004**, *34*, 112–124. [[CrossRef](#)] [[PubMed](#)]
13. Paul, M.J.; Meyer, J.L. Streams in the urban landscape. *Annu. Rev. Ecol. Syst.* **2001**, *32*, 333–365. [[CrossRef](#)]
14. Medupin, C.; Bark, R.; Owusu, K. Land Cover and Water Quality Patterns in an Urban River: A Case Study of River Medlock, Greater Manchester, UK. *Water* **2020**, *12*, 848. [[CrossRef](#)]
15. Astaraie-Imani, M.; Kapelan, Z.; Fu, G.; Butler, D. Assessing the combined effects of urbanisation and climate change on the river water quality in an integrated urban wastewater system in the UK. *J. Environ. Manag.* **2012**, *112*, 1–9. [[CrossRef](#)] [[PubMed](#)]
16. Hepp, L.U.; Milesi, S.V.; Biasi, C.; Restello, R.M. Effects of agricultural and urban impacts on macroinvertebrates assemblages in streams (Rio Grande do Sul, Brazil). *Zoologia* **2010**, *27*, 106–113. [[CrossRef](#)]
17. Moore, A.A.; Palmer, M.A. Invertebrate Biodiversity in Agricultural and Urban Headwater Streams: Implications for Conservation and Management. *Ecol. Appl.* **2011**, *15*, 1169–1177. [[CrossRef](#)]
18. Ongley, E.D.; Xiaolan, Z.; Tao, Y. Current status of agricultural and rural non-point source Pollution assessment in China. *Environ. Pollut.* **2010**, *158*, 1159–1168. [[CrossRef](#)]
19. Ortiz, J.D.; Puig, M.A. Point source effects on density, biomass and diversity of benthic macroinvertebrates in a mediterranean stream. *River Res. Appl.* **2007**, *23*, 16. [[CrossRef](#)]
20. UNFCC. *Climate Change and Freshwater Resources: A Synthesis of Adaptation Actions Undertaken by Nairobi Work Programme Partner Organisations*; UNFCC: Nairobi, Kenya, 2011.
21. Jiménez Cisneros, B.E.; Oki, T.; Arnell, N.W.; Benito, G.; Cogley, J.G.; Döll, P.; Jiang, T.; Mwakalila, S.S.; Kundzewicz, Z.; Nishijima, A. Freshwater resources. In *Climate Change 2014—Impacts, Adaptation and Vulnerability: Part A: Global and Sectoral Aspects*; Cambridge University Press: Cambridge, UK, 2015; pp. 229–270.
22. Jun, Y.C.; Kim, N.Y.; Kim, S.H.; Park, Y.S.; Kong, D.S.; Hwang, S.J. Spatial distribution of benthic macroinvertebrate assemblages in relation to environmental variables in Korean nationwide streams. *Water* **2016**, *8*, 27. [[CrossRef](#)]
23. Medupin, C. Distribution of benthic macroinvertebrate communities and assessment of water quality in a small UK river catchment. *SN Appl. Sci.* **2019**, *1*, 544. [[CrossRef](#)]
24. Melo, A.S. Explaining dissimilarities in macroinvertebrate assemblages among stream sites using environmental variables. *Zoologia (Curitiba)* **2009**, *26*, 79–84. [[CrossRef](#)]
25. Manzo, G. Complex Social Networks are Missing in the Dominant COVID-19 Epidemic Models. *Sociologica* **2020**, *14*, 31–49.
26. Wey, T.; Blumstein, D.T.; Shen, W.; Jordán, F. Social network analysis of animal behaviour: A promising tool for the study of sociality. *Anim. Behav.* **2008**, *75*, 333–344. [[CrossRef](#)]
27. Newman, M.E.J. Spread of epidemic disease on networks. *Phys. Rev. E* **2002**, *66*, 016128. [[CrossRef](#)] [[PubMed](#)]
28. Soul, J.; Hardingham, T.E.; Boot-Handford, R.P.; Schwartz, J.M. PhenomeExpress: A refined network analysis of expression datasets by inclusion of known disease phenotypes. *Sci. Rep.* **2015**, *5*, 8117. [[CrossRef](#)] [[PubMed](#)]
29. Su, G.; Morris, J.H.; Demchak, B.; Bader, G.D. Biological Network Exploration with Cytoscape 3. *Curr. Potocol Bioinform.* **2015**, *47*, 8–13. [[CrossRef](#)] [[PubMed](#)]
30. Yamanishi, Y.; Vert, J.P.; Kanehisa, M. Protein network inference from multiple genomic data: A supervised approach. *Bioinformatics* **2004**, *20*, 363–370. [[CrossRef](#)]

31. Hevey, D. Network analysis: A brief overview and tutorial. *Health Psychol. Behav. Med.* **2018**, *6*, 301–328. [[CrossRef](#)]
32. Clifton, A.; Webster, G.D. An Introduction to Social Network Analysis for Personality and Social Psychologists. *Soc. Psychol. Personal. Sci.* **2017**, *8*, 442–453. [[CrossRef](#)]
33. Schmittmann, V.D.; Cramer, A.O.J.; Waldorp, L.J.; Epskamp, S.; Kievit, R.A.; Borsboom, D. Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas Psychol.* **2013**, *31*, 43–53. [[CrossRef](#)]
34. Salpeteur, M.; Calvet-Mir, L.; Diaz-Reviriego, I.; Reyes-García, V. Networking the environment: Social network analysis in environmental management and local ecological knowledge studies. *Ecol. Soc.* **2017**, *22*. [[CrossRef](#)]
35. Allison, S.D.; Martiny, J.B.H. Resistance, resilience, and redundancy in microbial communities. *Light Evol.* **2009**, *2*, 149–166. [[CrossRef](#)] [[PubMed](#)]
36. Shade, A.; Read, J.S.; Youngblut, N.D.; Fierer, N.; Knight, R.; Kratz, T.K.; Lottig, N.R.; Roden, E.E.; Stanley, E.H.; Stombaugh, J.; et al. Lake microbial communities are resilient after a whole-ecosystem disturbance. *ISME J.* **2012**, *6*, 2153–2167. [[CrossRef](#)] [[PubMed](#)]
37. Ramirez, K.S.; Geisen, S.; Morriën, E.; Snoek, B.L.; van der Putten, W.H. Network Analyses Can Advance Above-Belowground Ecology. *Trends Plant Sci.* **2018**, *23*, 759–768. [[CrossRef](#)] [[PubMed](#)]
38. Frost, S.; Chiu, M.T.L.; Pugh Thomas, M. Seasonal changes of invertebrate populations in the polluted River Medlock. *Environ. Pollut.* **1976**, *11*, 223–242. [[CrossRef](#)]
39. James, P.; Atkinson, S.; Barlow, D.; Bates, A.; Comyn, F.; Duddy, M.; Dutton, D.; Fraser, J.; Horsfall, W.; Hothersall, A.; et al. *The Irwell Catchment Pilot: The Rivers Return*; The Environment Agency: Warrington, UK, 2012.
40. Council of the European Union. Water Framework Directive 2000/60/EC. *Off. J. Eur. Communities* **2000**, *72*. Available online: <https://eur-lex.europa.eu/eli/dir/2000/60/oj> (accessed on 15 September 2020).
41. APHA. *Standard Methods for the Examination of Water and Wate Water*; APHA: Washington, DC, USA, 1999.
42. WFD UK Technical Advisory Group. *UKTAG River Assessment Methods Benthic Invertebrate Fauna; Scottish Acid Water Indicator Community (SAWIC)*; WFD: Edinburgh, UK, 2009; ISBN 978-1-906934-19-4.
43. Jenson, K.S.; Dominigue, O.J. Extracting Topographic Structure from Digital Elevation Data for Geographic Information System Analysis. *Photogramm. Eng. Remote Sens.* **1988**, *54*, 1593–1600.
44. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.; Wang, J.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13*, 2498–2504. [[CrossRef](#)]
45. Smoot, M.E.; Ono, K.; Ruschinski, J.; Wang, P.L.; Ideker, T. Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics* **2011**, *27*, 431–432. [[CrossRef](#)]
46. Clarke, K.R.; Warwick, R.M. *Change in Marine Communities: An Approach to Statistical Analysis and Interpretation*, 2nd ed.; PRIMER-E Ltd. Plymouth Marine Laboratory: Plymouth, UK, 2001.
47. European Union. Water Framework Directive (WFD\_UKTAG) UKTAG WFD Environmental Standards River Basin Management 2015–2021. *Wfd\_Uktag* **2013**. Available online: <http://www.wfduk.org/sites/default/files/Media/Environmental%20standards/UKTAG%20Environmental%20Standards%20Phase%203%20Final%20Report%2004112013.pdf> (accessed on 15 September 2020).
48. Miserendino, M.L. Macroinvertebrate functional organization and water quality in a large arid river from Patagonia (Argentina). *Ann. Limnol.* **2007**, *43*, 133–145. [[CrossRef](#)]
49. Rawer-Jost, C.; Böhmer, J.; Blank, J.; Rahmann, H. Macroinvertebrate functional feeding group methods in ecological assessment. *Hydrobiologia* **2000**, *422–423*, 225–232. [[CrossRef](#)]
50. Medupin, C. Spatial and temporal variation of benthic macroinvertebrate communities along an urban river in Greater Manchester, UK. *Environ. Monit. Assess.* **2020**, *192*, 1–20. [[CrossRef](#)]
51. Harrell, F.E. Package ‘Hmisc’ 2020, 433. Available online: <https://github.com/harrelfe/Hmisc/> (accessed on 15 September 2020).
52. Leunda, P.M.; Oscoz, J.; Miranda, R.; Ariño, A.H. Longitudinal and seasonal variation of the benthic macroinvertebrate community and biotic indices in an undisturbed Pyrenean river. *Ecol. Indic.* **2009**, *9*, 53–63. [[CrossRef](#)]
53. Nelson, S.M. Response of stream macroinvertebrate assemblages to erosion control structures in a wastewater dominated urban stream in the southwestern U.S. *Hydrobiologia* **2011**, *663*, 51–69. [[CrossRef](#)]



54. Windsor, F.M.; Pereira, G.; Tyler, C.R.; Ormerod, S.J. Persistent contaminants as potential constraints on the recovery of urban river food webs from gross pollution. *Water Res.* **2019**, *163*, 114858. [[CrossRef](#)]
55. Jun, Y.-C.; Kim, N.-Y.; Kwon, S.-J.; Han, S.-C.; Hwang, I.-C.; Park, J.-H.; Won, D.-H.; Byun, M.-S.; Kong, H.-Y.; Lee, J.-E.; et al. Effects of land use on benthic macroinvertebrate communities: Comparison of two mountain streams in Korea. *Ann. Limnol. Int. J. Limnol.* **2011**, *47*, S35–S49. [[CrossRef](#)]
56. Uwadiae, R.E. Macroinvertebrates functional feeding groups as indices of biological assessment in a tropical aquatic ecosystem: Implications for ecosystem functions. *NY Sci. J.* **2010**, *3*, 6–15.
57. Xu, M.; Wang, Z.; Duan, X.; Pan, B. Effects of pollution on macroinvertebrates and water quality bio-assessment. *Hydrobiologia* **2014**, *729*, 247–259. [[CrossRef](#)]
58. Merritt, R.W.; Cummins, K.W. *An Introduction to the Aquatic Insects of North America*, 3rd ed.; Kendall/Hunt Publishing: Dubuque, IA, USA, 1996.
59. Palmer, C.G.; Maart, B.; Palmer, A.R.; O’Keeffe, J.H. An assessment of macroinvertebrate functional feeding groups as water quality indicators in the Buffalo River, eastern Cape Province, South Africa. *Hydrobiologia* **1996**, *318*, 153–164. [[CrossRef](#)]
60. Delmas, E.; Besson, M.; Brice, M.H.; Burkle, L.A.; Dalla Riva, G.V.; Fortin, M.J.; Gravel, D.; Guimarães, P.R.; Hembry, D.H.; Newman, E.A.; et al. Analysing ecological networks of species interactions. *Biol. Rev.* **2019**, *94*, 16–36. [[CrossRef](#)]
61. Pocock, M.J.O.; Evans, D.M.; Fontaine, C.; Harvey, M.; Julliard, R.; McLaughlin, Ó.; Silvertown, J.; Tamaddon-Nezhad, A.; White, P.C.L.; Bohan, D.A. The Visualisation of Ecological Networks, and Their Use as a Tool for Engagement, Advocacy and Management. *Adv. Ecol. Res.* **2016**, *54*, 41–85.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).