This document contains supplementary methods for the analysis described in Meyer et al. “Effects of spatially heterogeneous lakeside development on nearshore biotic communities in a large, deep, oligotrophic lake (Lake Baikal, Siberia)”. Additional details regarding data collection, aggregation, quality control, and recommendation for reuse can be found in the “protocols.pdf” file on the Environmental Data Initiative (Meyer et al. 2020) as well as this analysis’s companion data descriptor (Meyer et al. 2021).

*Inverse distance weighted (IDW) population calculation*

We recognized that sewage indicator concentrations at each sampling location may be related to a sampling location’s spatial position relative to both the size and proximity of neighboring developed sites. Therefore, we created the inverse distance weighted (IDW) population metric to compress, into a single metric, information about human population size, density, and location along the shoreline as well as distance between developed sites and sampling locations. The IDW metric reflects the idea that sewage pollution should be positively related to increasing human density and inversely related with distance from densely populated areas (sensu Bendz et al., 2005). Additionally, Timoshkin et al. (2018) noted that sewage enters Baikal’s nearshore largely through groundwater, implying that locations with more directly adjacent shoreline development should experience higher sewage concentrations in the lake. Acknowledging that nearshore PPCP concentrations were likely positively proportional to a development’s shoreline length, we scaled a developed site’s population density by its shoreline length. This scaling represents population density that directly interfaces with the lake, thereby capturing the idea that sewage-associated pollutants, such as PPCPs (Karnjanapiboonwong et al. 2010) and nutrients (de Vries 1972), contributed by a large development but away from the shoreline or with a small interface with the lake can be removed via the soil matrix en route to the lake.

Our calculation of IDW population was done in five steps. First, we traced polygons and shorelines from satellite imagery for each developed site in Google Earth. Polygons were traced for the entire area of visible development (Figure 1). Similarly, shoreline traces only reflected shoreline length for which there was visible development (Figure 1). Second, polygon and line geometries were downloaded from Google Earth as a .kml file. Third, the .kml file was imported into the R statistical environment (R Core Team 2019) where, using the sf package (Pebesma, 2018), we calculated shoreline length, polygon area, and centroid location for each developed site. Fourth, we joined point locations of each sampling site with the spatial polygons to calculate the distance from each sampling location to each developed site’s centroid. Fifth, we calculated IDW population for each sampling location, using formula (1)

(1)

where *I* is the IDW population at sampling location *j*, *P* is the population at each of the three developed sites Listvyanka (LI), Bolshie Koty (BK), Bolshoe Goloustnoe (BGO), *A* is the area of a developed site in km2, *L* is the shoreline length at a developed site in km, and *D* is the distance from sampling site *j* to each developed site’s centroid in km. This formulation implies that all sampling locations are influenced by all three developed sites. Thus, the influence of an individual developed site on each sampling location is positively influenced by the size and spatial density of the population and its orientation toward the shoreline, and inversely proportional to a sampling location’s distance from each of the three developed sites.

*Nutrients*

Water samples for nutrient analyses were collected in 150 mL glass jars that had been washed with phosphate-free soap and rinsed three times with water from the sampling location. Samples were collected in duplicates and immediately frozen at -20°C until processing at the A.P. Vinogradov Institute of Geochemistry (Siberian Branch of the Russian Academy of Sciences, Irkutsk). Samples were not filtered prior to freezing, meaning that nitrogen and ammonium concentrations may potentially include intracellular nitrogen and overestimate nitrogenous forms in the water column.

For each water sample, nitrate, ammonium, and total phosphorus concentrations were measured. For ammonium (RD:52.24.383-2018 2018) and nitrate (RD:52.24.380-2017 2018) concentrations, samples were analyzed with a spectrophotometer. Total phosphorus concentration was measured with a spectrophotometer following the addition of persulfate (GOST:18309-2014 2016). Concentrations are reported in mg/L.

*Chlorophyll a*

Water samples were collected in 1.5 L plastic bottles from a depth of approximately 0.75 m. Within 12 h of collection, three subsamples (up to 150 mL each) were filtered through 25-mm diameter, 0.2 µm pore size nitrocellulose filters. Filters were then placed in a 35-mm petri dish and frozen in the dark until processing.

Chlorophyll samples were processed in a manner similar to that of Welschmeyer (1994). Nitrocellulose filters were ground in 10 mL of 90% HPLC-grade acetone, in which chlorophyll extraction was allowed to proceed overnight. Chlorophyll extract was then analyzed using a Turner Designs 10-AU fluorometer (Turner Design, Sunnyvale, CA) using an excitation wavelength of 436 nm and emission of 680 nm. 10-AU Secondary Solid Standard (P/N 10-AU-904) was used to calibrate fluorometer prior to samples being processed. Blank samples registered a raw fluorescence of approximately 0.1 FL units. Concentrations were calculated using formula 2

.

Detection limits are estimated to be approximately 0.02 mg/L. Concentrations are reported as mg/L.

*PPCPs*

Water samples for PPCP analysis were collected in 250 mL amber glass bottles that were rinsed with either methanol or acetone and then three times with sample water prior to collections. Following collection, samples were refrigerated and kept in the dark until solid phase extraction (SPE). Due to the complexity of collecting these samples in Siberia, we did not collect field or lab blanks, but rather elected to collect more field samples along a larger shoreline transect.

Within 12 h of collection, samples were filtered directly from the amber glass bottle using an in-line Teflon filter holder with glass microfiber GMF (1.0 µm pore size, WhatmanGrad 934-AH) in tandem with a solid phase extraction (SPE) cartridge (200 mg HLB, Waters Corporation, Milford, MA) connected to a 1-liter vacuum flask. Lab personnel wore gloves and face masks to minimize contamination. Prior to filtration, SPE cartridges were primed with at least 5 mL of either methanol or acetone and then washed with at least 5 mL of sample water. Rate of extraction was maintained at approximately 1 drop per second. Extraction proceeded until water could no longer pass through the SPE cartridge or until all collected water was filtered. Cartridges were stored in Whirlpacks at -20°C until analysis for 18 PPCP residues using liquid chromatography tandem mass spectrometry (LC-MS-MS) following methods of Lee et al. (2016) and D’Alessio et al (2018). Concentrations are reported in µg/L.

*Microplastics*

At each location, samples were collected in triplicate using 1.5 L clear plastic beverage bottles that were washed thoroughly with sample water before each collection. While we do not know exactly what kind of plastic used for the bottle, they were likely made of polyethylene terephthalate. Samples were collected by hand for each littoral site and with a metal bucket from aboard the ship for pelagic sites.

For processing, each sample was vacuum filtered on to a 47-mm diameter GF/F filter. During filtration, aluminum foil was used to cover the filtration funnel to prevent contamination from airborne microplastic particles. After filtration, filters were dried under vacuum pressure and then stored in 50-mm petri dishes. Following filtration of all three replicates, the filtrate was collected and then re-filtered through a GF/F filter as a control for contamination from the plastic vacuum funnel or potentially airborne microplastics.

Microplastic counting involved visual inspection of the entire GF/F in a similar manner to methods described in Hanvey et al. (2017). Visual enumeration was conducted under a stereo microscope with ~100x magnification, and microplastics were classified into one of three categories: fibers, fragments, or beads. For all categories, plastics were defined as observed objects with apparent artificial colors (e.g., neon blue), so as to not enumerate plastics potentially contributed from the sampling bottle itself. Fibers were defined as smooth, long plastics with consistent diameters and no visible internal structures. Fragments were defined as plastics with irregularly sharp or jagged edges. Beads were defined as spherical plastics. Although we did not measure microplastic size, this technique likely allowed us to reliably quantify microplastics as small as ~300 µm (Hanvey et al. 2017). During enumeration, GF/Fs remained covered in the petri dish to minimize potential for contamination from the air. Following enumeration of both experimental and control samples, fibers, fragments, and beads enumerated in the controls were subtracted from the experimental microplastic densities for each plastic type and from each replicate. One location (BK-1) had two control replicates, which were averaged for each plastic type and then subtracted from the experimental samples. Results are reported as the average number of microplastics/L.

*Benthic algal collection*

At each littoral site, we haphazardly selected three rocks representative of local substrate. A plastic stencil was used to define a surface area of each rock from which we scraped a standardized 14.5 cm2 patch of periphyton. Samples were preserved with Lugol’s solution and stored in plastic scintillation vials. Additional periphyton was collected in composite from each site for fatty acid and stable isotope analysis.

Periphyton taxonomic identification and enumeration was performed by subsampling 10 μL aliquots from each preserved sample. For all 10 μL aliquots, cells, filaments, and colonies were counted, for the entire subsample, until at least 300 cells were identified for a given sampling replicate. If the first aliquot contained less than 300 cells, we counted additional subsamples until we reached at least 300 cells in total. In instances when 300 cells were counted before finishing a subsample, we still counted the entire aliquot. Taxa were classified into broad categories consistent with Baikal algal taxonomy (Izhboldina 2007), using coarse groupings to capture general patterns in relative algal abundance. As a result, algal groups consisted of diatoms, the filamentous green algal genera, *Ulothrix* spp. and *Spirogyra* spp.,and the green algal Order Tetrasporales.

*Benthic invertebrate collection*

At each littoral site, three kick-net samples were collected for assessment of benthic community composition and abundance. Using a D-net, we collected macroinvertebrates by flipping over 1-3 rocks, and then sweeping five times in a left-to-right motion across approximately 1 m. After the series of sweeps, the catch was rinsed into a plastic bucket. For each replicate, bucket contents were concentrated using a 64-μm mesh and placed in glass jars with 40% ethanol (vodka; the only preservative available to us at the time) for preservation and refrigerated at 4°C aboard the research vessel. The 40% ethanol preservative was replaced with ~80% ethanol upon return to the lab within 24 to 48 hours, and samples were stored at ~4°C.

Separate collections were conducted for invertebrate fatty acid and stable isotope analyses. Invertebrates were collected using a D-net and by hand. Organisms collected by hand included amphipod species that were observed from the community composition D-net collections but not readily observed in the stable isotope and fatty acids D-net collections. Collected organisms were live-sorted, identified to species, and then frozen at -20°C at the field station. The samples were later transferred to the lab in the U.S. via a Dewar flask with dry ice.

Invertebrate taxonomic identification and enumeration were performed under a stereo microscope. All adult amphipods were identified to species according to Takhteev and Didorenko (2015), whereas juveniles were identified to genus. Mollusks were identified to the family level according to Sitnikova (2012). Leeches were enumerated at the subclass level, but were likely all from the family Glossiphoniidae based on size, depth of sampling locations, and invertebrate communities sampled (Kaygorodova 2012). Caddisflies were enumerated at the order level, although Baikal does contain over 14 species of caddisfly (Valuyskiy et al. 2020). Flatworms were enumerated at the phylum level. All isopods enumerated were from the family Asellidae. Aside from having limited time available to spend with Baikal taxonomists during our field campaign, our choice of taxonomic resolution ultimately was a result of relative abundance for each taxonomic group, where amphipods were the most abundant taxa and flatworms were among the least abundant taxa across all sites (Figure S1). All samples contained oligochaetes and polychaetes, but due to poor preservation, these taxa were not counted. Six samples of the 42 collected were not well-preserved and were excluded from further analyses, in order to reduce errors in identification. KD-1 and LI-1 were the only sites with 1 sample counted. BK-2 and KD-2 each had two samples counted.

*Stable isotope analysis*

Measurements of δ15N and δ13C were performed on an elemental analyzer-isotope ratio mass spectrometer (EA-IRMS; Finnigan DELTAplus XP, Thermo Scientific) at the Large Lakes Observatory, University of Minnesota Duluth. The EA-IRMS was calibrated against certified reference materials including L-glutamic acid (NIST SRM 8574), low organic soil and sorghum flour (standards B-2153 and B-2159 from Elemental Micro-analysis Ltd., Okehampton, UK) and in-house standards (acetanilide and caffeine). Replicate analyses of external standards showed a mean standard deviation of 0.06 ‰ and 0.09 ‰, for δ13C and δ15N, respectively.

*Fatty acid analysis*

After overnight chloroform extraction, samples underwent a chloroform-methanol extraction three times. To each sample, we added 1 mL cooled 100% methanol, 1 mL chloroform:methanol solution (2:1), and 0.8 mL 0.9% NaCl solution. Samples were inverted three times and sonicated on ice for 10 minutes. Next, samples were vortexed for 1 minute, and centrifuged for 5 minutes (3,000 rpm) at 4°C. Using a double pipette technique, the lower organic layer was removed and kept under nitrogen. After the third extraction, samples were evaporated under nitrogen flow, and resuspended in 1.5 mL chloroform and stored at -20°C overnight.

Once resuspended in chloroform, 1 mL of chloroform extract was transferred to a glass centrifuge tube with a glass syringe as well as an internal standard of 4 μL of 19-carbon fatty acid. Samples were then evaporated under nitrogen, and then 1 mL of toluene and 2 mL of 1% sulfuric acid-methanol was added. The vial was closed under nitrogen gas and then incubated in 50°C water bath for 16 hours. After incubation, samples were removed from the bath, allowed to reach room temperature and stored on ice. Next, we performed a potassium carbonate-hexane elution twice. To each sample, we added 2 mL of 2% potassium bicarbonate and 5 mL of 100% hexane, inverting the capped vial so as to mix the solution. Samples were centrifuged for 3 minutes (1,500 rpm) at 4°C. The upper hexane layer was then removed and placed in a vial to evaporate under nitrogen flow. Once almost evaporated, 1 mL of 100% hexane was added to the fatty acid extract, and the sample was stored in a glass amber autosampler vial for GC/MS quantification. GC/MS quantification was performed with a Shimadzu QP2020 GC/MS following Schram et al. (2018).

*Inferring food web structure*

In order to assess food web structure, we compared stable isotopes and fatty acids in periphyton resources with those in amphipods. Periphyton and each amphipod species’ stable isotope signatures were each measured in composite for a given sampling location. Because periphyton stable isotope samples were measured for the aggregate community, periphyton could only be used as a single potential resource for inferring amphipod diets. Consequently, our stable isotope data constrained our analysis to qualitatively comparing δ13C and δ15N stable isotopes between periphyton and amphipod grazers. In contrast, algal fatty acid signatures tend to be consistent for individual taxa, allowing us to use published Baikalian algal fatty acid profiles with consumer fatty acid profiles from our sampling campaign to evaluate taxon-specific, algae-amphipod trophic interactions. We used a Bayesian mixing model to infer a potential resource’s relative abundance in amphipods’ diets using the MixSIAR framework (Stock et al. 2018b; a). This modelling procedure involved three data inputs:

1. Consumer Fatty Acids: These data were collected from our sampling at Lake Baikal. Because *E. verrucosus* and *E. vittatus* were most abundant along our disturbance gradient, we focused this analysis on those species’ fatty acid profiles.
2. Source Fatty Acids: Because our collected fatty acid data considered periphyton species in composite, we used published Baikalian taxon-specific fatty acid profiles to define characteristic diatom (Shishlyannikov et al. 2018) and *Ulothrix* spp. (Osipova et al. 2009) signatures. We used *Draparnaldia* spp. from our collected samples, as *Draparnaldia* spp.’s large cell sizes allowed us to isolate colonies with minimal other taxa. We *a priori* assumed that amphipods likely did not consume large filamentous algal taxa with cells larger than their mouthparts, such as *Draparnaldia* spp.; yet, we included filamentous fatty acids in our model as potential resources in the event amphipods were detritivorous on decomposing *Draparnaldia* spp. or *Ulothrix* spp. Therefore, including filamentous taxa as potential resources enabled us to account for nutrition that could be incorporated into the food web by grazers switching from herbivory to detritivory.
3. Trophic Discrimination Factors (TDFs): TDFs reflect shifts in fatty acid concentrations between trophic levels. To the best of our knowledge, there are no published TDFs for Baikal amphipods’ fatty acids. Therefore, we used TDFs estimated for Antarctic marine amphipods (Schram et al. 2019), which were fed diets of a single algal resource, as a proxy for Baikal amphipod TDFs. To ensure TDF estimates were conservative across consumers and resources, we averaged each fatty acid’s TDF, such that a given fatty acid’s TDF was identical for each potential resource.

Each consumer, source, and TDF file was then used as an input to MixSIAR. The model building procedure used uninformed prior distributions for each resource, a chain length of 100,000 with 50,000 burn-in, thin of 50, and 3 Monte Carlo Markov Chains. Because TDFs for this analysis were based on marine taxa, we assessed posterior sensitivity to TDF variation by increasing TDF standard deviations by 5%, 10%, 25%, 50%, and 100%, and then re-running the model. Chain convergence was assessed with a Gelman-Rubin diagnostic, which was below 1.005 for all model runs. Model fit was assessed by calculating RMSE twice (sensu Tanentzap et al. 2017): first, between mean consumer actual and predicted fatty acid proportions; second, between sample actual fatty acid proportions and predicted fatty acid proportions. Although MixSIAR has demonstrated robustness to deviations in TDFs and complexity of mixed resource inputs (Guerrero and Rogers 2020), this sensitivity analysis was designed to exceed errors that can arise from differences in mixing model methodologies and prescribed error structures (Happel et al. 2021). Each iteration of the sensitivity analysis produced a similar posterior result as the original TDFs. The accompanying R script “07\_foodweb\_analysis.R” details the exact data aggregation and model construction procedures and can be accessed from the project’s Open Science Framework portal (Meyer et al. 2015).

*Statistical analyses*

To identify areas of wastewater pollution with sewage indicators (Objective 1), total phosphorus, nitrate, ammonium, microplastic abundance and density, and total PPCP concentration were log-transformed and regressed against log-transformed IDW population using a linear model. Analytically, log-transforming made sites comparable, as values spanned three orders of magnitude. Physically, we assumed that these sewage indicators were likely subject to exponential processes (e.g., mixing, diffusion), and log-transforming the data should linearize the relationships between predictor and response variables. In contrast, variables that we *a priori* considered to not likely be influenced by mixing processes – chlorophyll a and δ15N values in tissues – were not log-transformed but still regressed against log-transformed IDW population using a linear model. Microplastic total abundance and concentration were evaluated separately as a subset of samples (n = 3) contained different volumes, allowing us to analyze whether total microplastics as well as density of microplastics differed along our transect. Residuals were assessed for normality and homogeneity of variance.

To increase confidence that our observed sewage indicator patterns were not a product of a limited sample size, we also performed a permutational analysis to compare how our actual results compared to a randomly permuted dataset. This process involved randomly permuting sewage indicator variables, regressing the respective sewage indicator against IDW population, and then extracting the p- and R2 values for the model. This routine was repeated 5,000 times for each sewage indicator, so as to generate a distribution of p- and R2 values that could have been possible, given our observed data. We then compared models’ p- and R2 values generated from non-permuted data to those from permuted datasets. If indeed models generated from observed data were describing a non-random process, p- and R2 values should be located at the tail end of the permuted values’ distribution. To summarize our original p- and R2 values in the context of those from models with permuted datasets, we report the percent of p-values less than and R2 values greater than those from models generated from non-permuted datasets.

To assess if benthic community composition was associated with increasing sewage indicators (Objective 2), periphyton and macroinvertebrate abundance data were each analyzed with a consistent multivariate workflow. First, replicates were averaged to characterize algal and macroinvertebrate compositions, and taxonomic groups representing less than 1% of the inter-site community were removed from analysis, in order to reduce the influence of rare species on results. Second, community compositions for both periphyton and macroinvertebrates were visualized using non-metric multidimensional scaling (NMDS) with a Bray-Curtis similarity metric. Periphyton community compositions were calculated as relative proportions, whereas invertebrate abundances were grouped at the genus-level for amphipods and at higher taxonomic levels for other macroinvertebrates and then square-root transformed to minimize influence of more abundant taxa. Amphipods were kept at the genus level because their numerical and relative abundance markedly exceeded the abundance of other macroinvertebrates (Figure S1). Visual inspection of the NMDS plot suggested that sites generally tended to separate by increasing PPCP concentrations and IDW population (see Table S2). To test whether sites’ benthic communities significantly differed with increasing PPCP concentration and IDW population, we first used k-mediods, also known as Partitioning Around the Mediods (PAM; Kaufman and Rousseeuw 2005), clustering to identify an optimal number of groupings. For this process, we iterated through multiple numbers of clusters (i.e., 1 to 10) and calculated the within-group-sum-of-squares (wss; Figure S2) and average silhouette width (Figure S3). We identified the optimal number of groups when wss decreased most markedly and when silhouette width was greatest (i.e., the elbow method) (Johnson and Wichern 2007). To confirm the optimal number as determined by non-hierarchical PAM clustering, we also used Weighted Pair-Group Centroid Clustering (WPGMC; Figure S4) as a hierarchical approach (Sneath and Sokal 1973), which corrects for clusters that may not be strongly discriminated regardless of how many samples are assigned to a given cluster (Legendre and Legendre 2012). We then performed three permutational multivariate analyses of variance (PERMANOVA; Anderson 2001) with 999 permutations: the first where community compositions were responses to the groups identified through clustering, the second where community compositions were responses to the continuous IDW population, and the third where community compositions were responses to total PPCP concentrations. When significant differences were identified, post-hoc SIMPER analysis (Clarke 1993) was performed following the PERMANOVA to identify which taxa contributed to 85% of the cumulative variance that most influenced IDW grouping separation.

To assess whether benthic food webs restructured with increasing sewage indicator concentrations (Objective 3), fatty acid data were analyzed in a manner similar to periphyton and macroinvertebrate abundance data. First, species’ fatty acid profiles were visualized by performing NMDS with Bray-Curtis similarity for all organisms’ relative fatty acid abundance (Figure S5). This technique broadly demonstrated that, as expected, interspecific variation in fatty acid composition was greater than intraspecific variation. The same pattern was observed for all fatty acids quantified as well as solely essential fatty acids (EFAs; Figure S6). Together, these NMDS plots suggested that periphyton fatty acids differentiated based on sewage indicator concentrations, which was likely a reflection of differences in periphyton community composition (Taipale et al. 2013). Among all taxa and sites, the fatty acids 18:3ω3, 18:1ω9, and 20:5ω3 had among the highest coefficients of variation, enabling comparisons between sites. These fatty acids tend to be associated with filamentous green algae (i.e., 18:3ω3 and 18:1ω9) and diatoms (i.e., 20:5ω3). To increase the robustness of our analysis, we expanded our approach to include major fatty acids within each taxonomic group, including 18:2ω6 (abundant in green algae); 16:1ω7 and 14:0 (abundant in diatoms); and 16:0 (abundant in both green algae and diatoms) (Taipale et al. 2013). To evaluate how relative fatty acid abundance may relate to sewage pollution, we assessed patterns among these seven fatty acids with both multivariate and univariate approaches. Within a multivariate framework, we created two NMDS plots with Bray-Curtis similarity, one just with primary producer (Figure S7) and the other with macroinvertebrate (Figure S8) fatty acid profiles. Because multivariate patterns suggested fatty acid profiles may relate to sewage pollution, we regressed a filamentous:diatom fatty acid ratio (Equation 2)

(2)

against log-transformed PPCP concentrations as well as IDW population using a linear model. Additionally, we evaluated how three essential fatty acids (18:3ω3, 18:2ω6, and 20:5ω3), lipids thought to accumulate in biological systems, may differ in abundance across the sewage gradient. Therefore, we similarly regressed the ratio of against log-transformed PPCP concentrations as well as IDW population using a linear model. As with sewage indicators, we recognized that these regression analyses and the associated interpretations may be compromised by a limited sample size. To ensure the robustness of these trends, we performed a permutational analysis similar to sewage indicators, where p- and R2 values for models generated from observed data were compared to models generated from 5,000 permutations.

All analyses were conducted in the R statistical environment (R Core Team 2019), using the tidyverse (Wickham et al. 2019), factoextra (Kassambara and Mundt 2019), cluster (Maechler et al. 2019), pvclust (Suzuki et al. 2019), ggrepel (Slowikowski 2019), viridis (Garnier 2018), fs (Hester and Wickham 2019), spdplyr (Sumner 2019), janitor (Firke 2020), sf (Pebesma 2018), ggpubr (Kassambara 2019), ggtext (Wilke 2020), OpenStreetMap (Fellows and Stotz 2019), cowplot (Wilke 2019), ggspatial (Dunnington 2021), broom (Robinson and Hayes 2019), ggsn (Baquero 2019), MixSIAR (Stock et al. 2018b), and vegan (Oksanen et al. 2019) packages. All data, including .kml files used to calculate IDW metric, are publicly available from the Environmental Data Initiative repository (Meyer et al. 2020), and all R scripts are available from the GitHub repository of this project’s Open Science Framework account (Meyer et al. 2015).

References

Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecology **26**: 32–46. doi:10.1111/j.1442-9993.2001.01070.pp.x

Baquero, O. S. 2019. ggsn: North Symbols and Scale Bars for Maps Created with “ggplot2” or “ggmap,.”

Clarke, K. R. 1993. Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology **18**: 117–143. doi:https://doi.org/10.1111/j.1442-9993.1993.tb00438.x

D’Alessio, M., S. Onanong, D. D. Snow, and C. Ray. 2018. Occurrence and removal of pharmaceutical compounds and steroids at four wastewater treatment plants in Hawai’i and their environmental fate. Science of The Total Environment **631–632**: 1360–1370. doi:10.1016/j.scitotenv.2018.03.100

Dunnington, D. 2021. ggspatial: Spatial Data Framework for ggplot2,.

Fellows, I., and using the Jm. library by J. P. Stotz. 2019. OpenStreetMap: Access to Open Street Map Raster Images,.

Firke, S. 2020. janitor: Simple Tools for Examining and Cleaning Dirty Data,.

Garnier, S. 2018. viridis: Default Color Maps from “matplotlib,.”

GOST:18309-2014. 2016. Methods for determination of phosphorus-containing matters (with corrections) (Методы определения фосфорсодержащих веществ).

Guerrero, A. I., and T. L. Rogers. 2020. Evaluating the performance of the Bayesian mixing tool MixSIAR with fatty acid data for quantitative estimation of diet. Sci Rep **10**: 20780. doi:10.1038/s41598-020-77396-1

Hanvey, J. S., P. J. Lewis, J. L. Lavers, N. D. Crosbie, K. Pozo, and B. O. Clarke. 2017. A review of analytical techniques for quantifying microplastics in sediments. Anal. Methods **9**: 1369–1383. doi:10.1039/C6AY02707E

Happel, A., J. Pike, S. Czesny, and J. Rinchard. 2021. An empirical test of fatty acid based diet estimation models. Food Webs e00197. doi:10.1016/j.fooweb.2021.e00197

Hester, J., and H. Wickham. 2019. fs: Cross-Platform File System Operations Based on “libuv,.”

Izhboldina, L. A. 2007. Guide and Key to Benthic and Periphyton Algae of Lake Baikal (meio- and macrophytes) with Brief Notes on Their Ecology, Nauka-Centre.

Johnson, R. A., and D. V. Wichern. 2007. Applied Multivariate Statistical Analysis, 6th ed. Prentice Hall.

Karnjanapiboonwong, A., A. N. Morse, J. D. Maul, and T. A. Anderson. 2010. Sorption of estrogens, triclosan, and caffeine in a sandy loam and a silt loam soil. Journal of Soils and Sediments **10**: 1300–1307. doi:10.1007/s11368-010-0223-5

Kassambara, A. 2019. ggpubr: “ggplot2” Based Publication Ready Plots,.

Kassambara, A., and F. Mundt. 2019. factoextra: Extract and Visualize the Results of Multivariate Data Analyses,.

Kaufman, L., and P. J. Rousseeuw. 2005. Finding Groups in Data: An Introduction to Cluster Analysis, 1st Edition. Wiley-Interscience.

Kaygorodova, I. 2012. An Illustrated Checklist of Leech Species from Lake Baikal (Eastern Siberia, Russia). Dataset Papers in Biology **2013**: e261521. doi:10.7167/2013/261521

Legendre, P., and L. Legendre. 2012. Numerical Ecology, 3rd ed. Elsevier.

Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, and K. Hornik. 2019. cluster: Cluster Analysis Basics and Extensions,.

Meyer, M. F., T. Ozersky, K. H. Woo, and others. 2020. A unified dataset of co-located sewage pollution, periphyton, and benthic macroinvertebrate community and food web structure from Lake Baikal (Siberia).doi:10.6073/PASTA/76F43144015EC795679BAC508EFA044B

Meyer, M. F., T. Ozersky, K. H. Woo, and others. Under Revision. A unified dataset of co-located sewage pollution, periphyton, and benthic macroinvertebrate community and food web structure from Lake Baikal (Siberia).

Meyer, M., T. Ozersky, K. Woo, A. W. E. Galloway, M. R. Brousil, and S. Hampton. 2015. Baikal Food Webs.doi:10.17605/OSF.IO/9TA8Z

Oksanen, J., F. G. Blanchet, M. Friendly, and others. 2019. vegan: Community Ecology Package,.

Osipova, S., L. Dudareva, N. Bondarenko, A. Nasarova, N. Sokolova, L. Obolkina, O. Glyzina, and O. Timoshkin. 2009. Temporal variation in fatty acid composition of Ulothrix zonata (Chlorophyta) from ice and benthic communities of Lake Baikal. Phycologia **48**: 130–135.

Pebesma, E. 2018. Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal **10**: 439–446. doi:10.32614/RJ-2018-009

R Core Team. 2019. R: A Language and Environment for Statistical Computing,.

RD:52.24.380-2017. 2018. Nitrate concentration in waters: Photometric methods with Giress reagent following stabilization in a cadmium reducer (Массовая концентрация нитратного азота в водах: Методика измерений фотометрическим методом с реактивом Грисса после восстановления в камиевом редукторе).

RD:52.24.383-2018. 2018. Working Document: Concentration of aqueous ammonium: Method for measuring with a photometer using indophenol blue (Руководящий Документ: Массовая концентрация аммонийного азота в водах: Методика измерений фотометрическим методом в виде индофенолового сингео). RD:52.24.383-2018. RD:52.24.383-2018.

Robinson, D., and A. Hayes. 2019. broom: Convert Statistical Analysis Objects into Tidy Tibbles,.

Schram, J. B., M. O. Amsler, A. W. E. Galloway, C. D. Amsler, and J. B. McClintock. 2019. Fatty acid trophic transfer of Antarctic algae to a sympatric amphipod consumer. Antarctic Science **31**: 315–316. doi:10.1017/S0954102019000397

Schram, J. B., J. N. Kobelt, M. N. Dethier, and A. W. E. Galloway. 2018. Trophic Transfer of Macroalgal Fatty Acids in Two Urchin Species: Digestion, Egestion, and Tissue Building. Front. Ecol. Evol. **6**. doi:10.3389/fevo.2018.00083

Shishlyannikov, S. M., A. A. Nikonova, Y. S. Bukin, and A. G. Gorshkov. 2018. Fatty acid trophic markers in Lake Baikal phytoplankton: A comparison of endemic and cosmopolitan diatom-dominated phytoplankton assemblages. Ecological Indicators **85**: 878–886. doi:10.1016/j.ecolind.2017.11.052

Sitnikova, T. Ya. 2012. Определитель брюхоногих моллюсков бухты Большие Коты (юго-западное побережье озера Байкал) [Key of the Gastropod Molluscs in the Bay of Bolshie Koty (South-West shoreline of Lake Baikal)], Irkutsk State University.

Slowikowski, K. 2019. ggrepel: Automatically Position Non-Overlapping Text Labels with “ggplot2,.”

Sneath, P. H. A., and R. R. Sokal. 1973. Numerical Taxonomy: The Principles and Practice of Numerical Classification, W. H. Freeman.

Stock, B. C., A. L. Jackson, E. J. Ward, A. C. Parnell, D. L. Phillips, and B. X. Semmens. 2018a. Analyzing mixing systems using a new generation of Bayesian tracer mixing models. PeerJ **6**: e5096. doi:10.7717/peerj.5096

Stock, B., A. Jackson, E. Ward, and J. Venkiteswaran. 2018b. brianstock/MixSIAR 3.1.9, Zenodo.

Sumner, M. D. 2019. spdplyr: Data Manipulation Verbs for the Spatial Classes,.

Suzuki, R., Y. Terada, and H. Shimodaira. 2019. pvclust: Hierarchical Clustering with P-Values via Multiscale Bootstrap Resampling,.

Taipale, S., U. Strandberg, E. Peltomaa, A. W. E. Galloway, A. Ojala, and M. T. Brett. 2013. Fatty acid composition as biomarkers of freshwater microalgae: analysis of 37 strains of microalgae in 22 genera and in seven classes. Aquatic Microbial Ecology **71**: 165–178. doi:10.3354/ame01671

Takhteev, V. V., and D. I. Didorenko. 2015. Fauna and ecology of amphipods of Lake Baikal: A Training manual, V.B. Sochava Institute of Geography SB RAS.

Tanentzap, A. J., B. W. Kielstra, G. M. Wilkinson, and others. 2017. Terrestrial support of lake food webs: Synthesis reveals controls over cross-ecosystem resource use. Science Advances. doi:10.1126/sciadv.1601765

Valuyskiy, M. Y., S. I. Melnitsky, and V. D. Ivanov. 2020. Structure and Evolution of the Antennal Sensory Surface in Endemic Caddisfly Tribes Baicalinini and Thamastini (Trichoptera: Apataniidae) from Lake Baikal. J Evol Biochem Phys **56**: 318–332. doi:10.1134/S0022093020040031

de Vries, J. 1972. Soil Filtration of Wastewater Effluent and the Mechanism of Pore Clogging. Journal (Water Pollution Control Federation) **44**: 565–573.

Welschmeyer, N. A. 1994. Fluorometric analysis of chlorophyll a in the presence of chlorophyll b and pheopigments. Limnol. Oceanogr. **39**: 1985–1992. doi:10.4319/lo.1994.39.8.1985

Wickham, H., M. Averick, J. Bryan, and others. 2019. Welcome to the tidyverse. Journal of Open Source Software **4**: 1686. doi:10.21105/joss.01686

Wilke, C. O. 2019. cowplot: Streamlined Plot Theme and Plot Annotations for “ggplot2,.”

Wilke, C. O. 2020. ggtext: Improved Text Rendering Support for “ggplot2,.”