HOW MUCH SUBSAMPLING IS ENOUGH?: THE DEVELOPMENT OF A STATISTICALLY JUSTIFIED METHOD FOR SAMPLING FILTERED MICROPLASTICS IN DRINKING WATER

by

Elysia Fuller-Thomson

A Thesis
Submitted to the School of Interdisciplinary Science
In Partial Fulfilment of the Requirements
For the Degree
Bachelor of Science

McMaster University
© Copyright by Elysia Fuller-Thomson, 2021

BACHELOR OF SCIENCE (2021) McMaster University

School of Interdisciplinary Science Hamilton, Ontario

TITLE: How much Sampling Area is Enough?: The Development of a Statistically Justified Method for Sampling Filtered Microplastics in Drinking Water

AUTHOR: Elysia Fuller-Thomson

SUPERVISOR: Dr. Robert Andrews, Dr. Michael McKie and Dr. Carolyn Eyles

NUMBER OF PAGES: vi, 33

Abstract

As awareness intensifies regarding the presence of microplastics (MPs) throughout the environment, there is a growing concern that MPs in drinking water may contribute to a toxic response in consumers. With the increasingly large volumes (> 100L) being sampled to improve accuracy, yielding potentially hundreds of MPs that must be counted, coupled with the large amount of time it requires to process MP samples, there is a need to know the minimum subsample required to accurately represent the actual MP population. In order to develop a statistically justified framework for sampling, drinking water was filtered via “in-lab” and “in-line” methods. Fifteen 10 μm filters 47 mm in diameter were examined with 70 0.7 mm^2 grid cells, where each cell’s total MPs were counted using an optical microscope. Suspected MPs were examined in areas ranging within zero to 100% of the filter using three strategies: i) random discontiguous, ii) one random contiguous area, and iii) two random contiguous areas equidistant from the centre. Significant clustering was observed in all but two of the fifteen filters. Subsample estimates of total suspected MPs were compared with the actual MP population via consistency values. A minimum suitable sampling area was determined to exist when a specified subsample area and every sampling interval above it had a 95% confidence interval of each filter’s upper and lower confidence intervals of consistency that fell within a suitable percentage from the true value. Results of this study suggest that out of all the sampling methods tested, discontiguous sampling is optimal. Randomly sampling one quarter of the filter results in an accuracy within ± 15%, nineteen times out of twenty. Overall, it is recommended that this method be employed for future research in order to statistically justify the sampling area required for MPs in drinking water.

Acknowledgements

I would like to thank Dr. Robert Andrews for his thoughtful feedback throughout the thesis process. I would also like to thank Dr. Michael McKie for his continual support and insightful advice from the initial proposal to the final submission. Thanks to Chuqiao Yuan for so kindly providing me with her data and answering my many questions. Thank you to Dr. Carolyn Eyles for having the confidence in me to allow my thesis work to be pursued outside of McMaster University.

A special thanks to Mum, Dad, Justin, Lilia and Ian for modeling what it means to have a heart for understanding the workings of the world around us and a never-ending love of learning.

Contents

List of Figures

List of Tables

Nomenclature

%	Percent
±	Plus minus
<	Less than
≤	Less than or equal to
>	Greater than
≥	Greater than or equal to
~	Approximately
In-lab	In-laboratory
L	Litre
Min.	Minimum
mm	Millimetre
MP	Microplastic
µm	Micrometre
QA/QC	Quality assurance/quality control

Introduction

Presence of Microplastics

Microplastics (MPs) have been detected in freshwater and saltwater ecosystems (Auta, Emenike and Fauziah, 2017; Li, Liu and Paul Chen, 2018), the atmosphere (Chen, Feng and Wang, 2020), and terrestrial environments around the world (Qi et al., 2020). The California State Water Resources Control Board has recently defined MPs as manmade solid polymeric particles with dimensions between 1 nm and 5 mm in size (Adoption of Definition of Microplastics in Drinking Water), which accurately describes the definition agreed upon within the literature (Koelmans et al., 2019; Novotna et al., 2019). As MPs are considered an emerging contaminant and potential human health hazard, concern has grown regarding the presence of MPs in drinking water (WHO, 2019). MPs have been quantified throughout the drinking water treatment process including untreated, raw water and treated tap water (Novotna et al., 2019). While there is currently no legislation limiting the concentration MPs within treated water, it is likely that existing technology within drinking water treatment plants is capable of removing at least >70 % of MPs present in source waters (Pivokonsky et al., 2018a; Wang, Lin and Chen, 2020a).

It is currently estimated that drinking water contributes to approximately 1% of all the MPs that are consumed by humans (Van Cauwenberghe and Janssen, 2014; Cox et al., 2019). Although MPs have been consistently observed in raw and treated water, their concentration ranges widely between studies, from <1 MP/L to >1000 of MPs/L (Mintenig et al., 2019; Wang, Lin and Chen, 2020; Pivokonsky et al., 2018; Novotna et al., 2019). While much of this difference is related to source water concentrations, differing sampling and counting methods, sample contamination as well as inconsistent quality assurance (QA) and quality control controls (QC) make it challenging to compare reported MPs among studies (Koelmans et al., 2019). Improving the comparability of

studies by improving sampling procedures may serve as a major next step in conducting MP risk assessment.

Health Concerns

While consistent evidence exists for MP toxicity in aquatic organisms, the evidence for human toxicity has not yet well been elucidated (WHO, 2019). MPs may contribute to potential human health issues as a function of: i) particle physical toxicity, ii) chemical toxicity and, iii) as biofilm hosts (WHO, 2019). Deng et al. (2017) reported that MP ingestion resulted in physical toxicity in mice after 4 weeks of exposure to 0.1 mg MPs/day MPs. The mice were reported to experience significant decreases in ATP levels and relative liver weight. Key neurotransmitter precursors also decreased, an indication of potential MP derived neurotoxicity. Lu et al. (2018) observed gut microbiome disruptions when exposed to 1.456×10^{10} particles/L 0.5 μm in size as well as 1.456×10^{4} particles/L 50 μm in size for five weeks. In human cells, MP exposure of 1000 μg/L has resulted in responses that induce inflammation (Hwang et al., 2020). These results imply that consumption of MPs by humans may physically incur a toxic response and reinforces the need to accurately evaluate the quantity and type of MPs in treated drinking water.

While the main components of MPs are chemically inert polymers, up to 4% of MPs may include potentially harmful residual agents, including plasticizers such as phthalates and flame retardants such as polybrominated diphenyl ethers (WHO, 2019). These agents sometimes remain following chemical processing and induce chemical toxicity (Rist et al., 2018). Bisphenol A, an endocrine disruptor, is a widely known example of these chemicals, however many thousands more exist (Rist et al., 2018). Further adding to potential toxicity pathways, sorbed hydrophobic chemicals can bind to MPs and cause harm (Rist et al., 2018; Hartmann et al., 2017). For example, Avio et al. (2015) found that after six days of treatment of $\geq$0.5 μg/L pyrene, polyethylene and polystyrene

MPs absorbed 145 ± 35 ng/g and 126 ± 35 ng/g of pyrene, respectively. Pittura et al. (2018) likewise reported benzo(a)pyrene, a polyaromatic hydrocarbon, sorbed to MPs. Upon ingestion by mussels, there was an induced cellular toxicity after four weeks of exposure to 10 mg/L of benzo(a)pyrene-contaminated MPs which presented evidence for the possible transfer of benzo(a)pyrene from MPs to the organism. In surface water environments, MPs have also been identified as vectors for human bacterial pathogens, especially *Vibrio* species (Bowley et al., 2020).

Unfortunately, studies regarding the impact of MPs on human health studies are absent; laboratory trials to-date provide insufficient evidence to make definitive conclusions on MP human health concerns (WHO, 2019). Human health hazards may be defined as a product of both toxicity and exposure (WHO, 2019). As toxicity research is ongoing, research on exposure (e.g., concentration and composition) through drinking water is an essential part of assessing overall MP risk.

Microplastic Sample Analysis and Characterization

To better understand MPs in drinking water systems, water samples have been collected from water treatment facilities (Johnson et al., 2020), wells which serve as sources to drinking water systems (Panno et al., 2019), and from consumer taps (Shruti, Pérez-Guevara and Kutralam-Muniasamy, 2020). The literature varies with respect to the need of replicates with studies ranging from individual samples to 3-10 replicates (Table 1; Pivokonsky et al., 2018; Wang, Lin and Chen, 2020). Samples may be subsequently stained, centrifuged or mixed with a surfactant to enhance MP detection and quantification (Elkhatib and Oyanedel-Craver, 2020; Cherniak et al., 2020). Separation of suspected MPs from water is typically accomplished using either a stack of sieves of varying mesh sizes or membrane filters ranging from 5 to 5000 μm (Elkhatib and Oyanedel-Craver, 2020; Johnson et al., 2020). Due to low concentrations commonly observed in many

treated drinking waters, high-throughput in-line filtration consisting of cartridge filters within a filter housing may be employed such that larger volumes of water (>100L) may be evaluated without the need to transport exceedingly large volumes of water to an analytical laboratory (Mintenig et al., 2019). Larger volumes are expected to both provide more representative results and ensure that a sufficient number of MPs are evaluated.

Following filtration, the filter is typically scanned visually for suspected MPs using microscopy such that they may be classified according to physical characteristics, including size, shape (e.g. fibers, films, fragments) and colour (Shruti, Pérez-Guevara and Kutralam-Muniasamy, 2020). This process can be accomplished manually or through the use of visual analysis software such as ImageJ (Rivers, Gwinnett and Woodall, 2019). Suspected MPs must then be characterized as to their chemical composition using Fourier transform infrared (FTIR) or Raman Spectroscopy (Silva et al., 2018). Some studies have combined physical and chemical quantification via FTIR micro spectroscopy using complementary software (Mintenig et al., 2019; Johnson et al., 2020).

In decreasing order, the five most common shapes represented in the literature are fragment, fibres, films, foams and pellets (Koelmans et al., 2019). In general, MPs of smaller sizes appear far more abundantly. According to Pivokonsky et al. (2018), 95% of particles in raw and treated drinking water have been reported with a size <10 μm. Reflecting trends in global plastic production, the three most common plastic types, from highest to lowest presence, are polyethylene, polypropylene, and polystyrene (Elkhatib and Oyanedel-Craver, 2020). Plastics with densities greater than water (>1 g/ml) tend to settle out of aquatic environments and are not observed in as high proportions relative to their abundance in global production (Koelmans et al., 2019).

Current Drinking Water Microplastic Methodology Recommendations

Koelmans et al. (2019) presented a systematic review of sampling procedures which highlighted key issues regarding methodology for the determination of MP concentration. They suggested a number of criteria for improving MP sampling strategies including the use of a minimum sample volume of 500 L and 1000 L for surface and treated water, respectively. These volumes were suggested to provide a representative sample, defined as the analysis of 5 to 500 MPs. The same authors suggest using FTIR or Raman spectroscopy on a minimum of 50 particles per sample or 50% of the total particle number. Furthermore, they suggest at least 25% of the filter should be observed (Koelmans et al., 2019). While Koelman's et al. (2019) suggests these methodologies to improve accuracy, sampling procedures reported in the literature vary widely and are rarely as rigorous. As well, it should be noted that differences in source water, seasonality, and applied treatment technologies may limit the transferability of the predetermined sample sizes to new situations, often requiring pre-sampling to ensure that appropriate volumes are collected based on the MP concentration (Wang et al., 2021; Johnson et al., 2020). Many authors highlight the concern of the time intensive nature of MP chemical verification via Raman spectroscopy, showing that well thought-out, timely subsampling for qualitative analysis is needed (Pivokonsky et al., 2018; Johnson et al., 2020).

In the literature, methods vary widely, as well as their final MP concentration estimates. Table 1 highlights just how unstandardized methodology remains in drinking water to this day. Volumes sampled range several orders of magnitude. Filter area sampled ranges from <1 % of the filter to 100% of the filter. Minimum particle size varies from <1 μm to 100 μm. As such, papers that report smaller particle size will likely observe higher concentrations of MPs solely because they include MPs that were too small to detect in other studies (Johnson et al., 2020). With such large

Table 1: Methods of selected drinking water sampling and analysis from the literature.

Author	Source Water	Volume of Water Sampled per Sample (L)	Minimum Particle Size (μm)	Number of Samples Taken per Exact Location	Percentage of Filter Area Chosen for Quantitative Analysis	Percentage of Filter Area Chosen for Qualitative Analysis	Average MP Concentration in Raw Water (MP/L)	Average MP Concentration in Treated or Tap Water (MP/L)
Adib, Mafigholami and Tabeshkia, 2021	Surface	2.5	1	3	.7%	107 for all filters	2325±101	1138±100
Pittroff et al., 2021	Ground	> 1300	5	1	22%[a]	22%[a]	0.066 ± 0.076[b]	0.066 ± 0.076[b]
Sarkar et al., 2021	Surface	30	25	15	100%	100%	17.88	2.75
Cherniak et al., 2020	Surface	~10	10	2	100%	10% of particles (min. 10 MPs)	42 ± 18	20 ± 8
Johnson et al., 2020	Surface & Ground	> 100	5	5	92 %	92 %	4.9	0.0011
Mintenig et al., 2019	Ground	300 - 2500	3	1	100%	100 %	0.0007[b]	0.0007[b]
Shruti, Pérez-Guevara and Kutralam-Muniasamy, 2020	N/A	1	.22	1	100%	Unreported	N/A	18 ± 7
Tong et al., 2020	N/A	2[b]	0.2	1	0.66%	A Cross Section of Both Axes	N/A	440 ± 275
Pivokonský et al., 2020	Surface	2[c]	1	3	54%	25%	23 ± 2[d] 1296 ± 35	14 ± 1[d] 151 ± 4
Panno et al., 2019	Ground	2	0.45	1	100%	20 MP per filter	2.76	N/A
Pivokonský et al., 2018	Surface	1[e]	1	9	15%	25%	2297 ± 166	470 ± 27
Strand et al., 2018	N/A	50	100	1	100%	10%	N/A	0.58

Notes:
Only studies that included some chemical identification and evidence of MPs as well as physical quantification are included.
[a]For particles less than 10 μm, only 1.4% of the filter area was analyzed.
[b]Raw and treated water concentrations were not reported separately, because differences within treatment were far more apparent than between treatment.
[c]1 L was collected for quantitative analysis + 1 L for qualitative analysis.
[d]The DWTPs are disclosed separately due to their extreme differences
[e]1 L were collected for quantitative analysis per replicate and 1 L per sampling day (three days total) was taken for qualitative analysis

discrepancy in methodology within the literature, it is difficult to distinguish whether the large observed differences in concentration are due to the intrinsic properties of the source water or an outcome of methodology dissimilarities. Study after study flags the need to standardize methodology in order to alleviate this concern (Danopoulos, Twiddy and Rotchell, 2020; Elkhatib and Oyanedel-Craver, 2020; Koelmans et al., 2019; Zhang et al., 2020).

Challenges in Accurately Quantifying Microplastics

It remains difficult to quantify and characterize MPs in drinking waters due to analytical challenges. Contamination remains an ever-present concern as laboratory environments (e.g. air and laboratory equipment) may contain or be comprised of plastic materials (Wesch et al., 2017). A further lack of positive and negative controls, including environmental spiking and blanks, along with insufficient chemical analysis, raises issues regarding the confidence of published MPs concentrations (Koelmans et al., 2019).When considering the relatively small number of MPs typically observed in treated drinking water, small volumes of water (<25 L) may not be adequate for accurate MP concentration assessment (Koelmans et al., 2019). However, transferring and storing large volumes, while also protecting them from airborne contamination is not practical. In addition, the large amount of time required to visually characterize and chemically analyze suspected MPs significantly hinders the ability to quantify MPs in large numbers (> 100 MPs) (Pivokonsky et al., 2018; Wang, Lin and Chen, 2020).

Recent studies have drawn attention to ensuring clean laboratory environments and the analysis of positive and negative controls in an effort to minimize contamination (Cherniak et al., 2020; Johnson et al., 2020). To further reduce contamination during sample collection, high-throughput, in-line filtration is a relatively new form of sampling that allows for hundreds of liters of water to be sampled, while effectively reducing exposure to airborne contaminants (Mintenig et al., 2019).

Three publications of note have utilized high-throughput in-line filtration for MP quantification in raw and treated water (Johnson et al., 2020; Mintenig et al., 2019; Pittroff et al., 2021). Methodologically, the studies' detection of low MP concentrations in surface water suggests perhaps an even higher sampling volume is required than that suggested by Koelman et al. (2019). In-line filtration suggests much lower values for MP concentrations in raw and treated waters when compared to in-lab filtration, with treated water concentrations <0.1 MP/L. Values for treated water are several orders of magnitude below the predominant concentration within the literature, suggesting a potential lack of scalability when considering smaller sample volumes. Should high-throughput in-line MP quantification research continue to suggest radically different results when compared to low volume measurements, the implications are broad. Large volumes provide the potential to capture more MPs, however, more particles inherently require additional time to accurately identify and characterize. This issue is compounded when employing high-throughput in-line filtration as large sample volumes (> 1 m^3) could potentially yield hundreds of MPs.

Improved reporting of QA/QC controls and procedures have emerged following concerns regarding potential MP contamination (Wesch et al., 2017; Koelmans et al., 2019), however, the field needs to improve the justification of sampling procedures in order to close research gaps. Ideally, individual studies should always report volume of water filtered, number of replicates, and proportion of filter area selected for analysis. Hence there remains a need to the question: how many MPs do we need to sample? No studies could be identified that specifically justify exact volumes, number of replicates, or the amount of filter area selected for analysis. As such, there exists a need to develop a statistically justified approach for subsampling MPs collected from water treatment facilities in order to reduce analysis time without negatively impacting data quality. For the purposes of this study, a statistically justified sampling procedure is represented by a stepwise

process to provide confidence that the sample is representative of the whole population within a predetermined suitable range. Since subsample filter area is not directly related to source concentration, this parameter is one of the most general sampling decisions applied to almost all MP drinking water analysis. In order to improve MP sampling methodology with a particular focus on subsampling filter area, the main objectives of this study were to: i) develop a method to provide statistical justification for selecting the area of filter to analyze MPs and ii) to suggest an optimal value for the area of a filter that should be subsampled for drinking water samples.

Method

Sample Collection and Processing

In order to determine a statistically suitable sampling area within a given filter, suspected MPs were filtered, repeatedly sampled, and finally compared to the population counted on the entire filter. While filtration is described in greater detail by Yuan (2021) as well as Cherniak et al. (2020), the filters analyzed in this study considered samples obtained using two principal filtration strategies: i) in-line and ii) in-lab filtration (Table 2). Samples were collected from the influent and effluent flow of a water treatment plant which treats water originating from Lake Simcoe, Ontario. In-line filtration incorporates a 10 μm polycarbonate filter 47 mm in diameter (Millipore Sigma, Burlington, MA, USA) housed within a stainless-steel filter holder (Sigma-Aldrich, Oakville, ON, CA) followed by a nickel-plated brass flow totalizer (McMaster-Carr, Elmhurst, Il, USA).

In contrast, in-lab filtration required water samples to be collected in 20 L stainless steel Cornelius Ball Lock Kegs (Ontario Beer Kegs, Bolton, ON, CA), which were previously cleaned and rinsed three times with Elix water (MilliPore Sigma, Burlington, MA, USA) (Type II ASTM). After sample collection, the kegs were transported to the University of Toronto Drinking Water Research Group laboratory. Alcojet® surfactant was then added to achieve a 1% surfactant solution

Table 2: Properties of the source water and filtration strategy of the filters

Filter	Source Water	Filtration Strategy	Influent or Effluent	Volume (L)	Sample Number	Rinse/Clog
1	WTP Sample	In-line	Influent	189.2	1	
2	WTP Sample	In-line	Influent	<1	1	Rinse for Filter #1
3	WTP Sample	In-line	Influent	6.4	1	
4	WTP Sample	In-line	Influent	<1	1	Rinse for Filter #3
5	Field Blank (Elix Water)	In-lab	Influent	9	2	
6	Field Blank (Elix Water)	In-lab	Influent	9	2	
7	WTP Sample	In-lab	Effluent	12	2	
8	WTP Sample	In-lab	Effluent	12	1	Rinse for Filter #7
9	WTP Sample	In-lab	Effluent	13	2	
10	WTP Sample	In-lab	Influent	13	1	Rinse for Filter #9
11	WTP Sample	In-lab	Influent	12	2	
12	WTP Sample	In-lab	Influent	12	2	Second filter due to clog
13	WTP Sample	In-lab	Influent	12	1	Rinse for filter 11 & 12
14	WTP Sample	In-lab	Influent	14	2	
15	WTP Sample	In-lab	Influent	14	2	

(Alconox, White Plains, NY, USA). The kegs were pressurized (Mastercraft, Canadian Tire, Toronto, ON) with filtered air (Cole Parmer, Montreal, QC) to flush the contents through copper tubing and a 10 μm pore filter 47 mm in diameter. The kegs were subsequently rinsed three times using 1% Alcojet® solution and then pressurized to ensure that all of the sample was expelled. Prior to use, all surfaces, materials, and instruments to potentially come into contact with the sample were cleaned and rinsed three times with Elix water. A HEPA air purifier (ALEN, Austin,

TX, USA) was in use for the entirety of the sampling and counting process to minimize air contamination.

Counting and Characterization

Following the method reported by Yuan (2021), a 10 μm pore size filter was placed above a Petri-Sticker grid of 70 cells with dimensions of 7 mm by 7 mm (Diversified BiotechSo, Dedham, MA, USA) on top of a clear glass petri dish (VWR, Mississauga, ON, Canada). Under an OMAX 10X - 80X binocular zoom stereo microscope (Microscopenet.com, Kitchener, ON, Canada), each grid was counted for MP suspected particles and their colour and shape (e.g. fragment, fibre, or film) were recorded.

Locating Suspected MPs on Filters

Since the MP count had been used for a previous study, identification of the outside perimeter of the filters on the numbered grids had to be estimated. To do this, the filter was located using these three criteria of decreasing importance: i) no part of the filter was located outside the numbered cells of the PetriSticker, ii) every cell with a recorded suspected MP particle must be located in a numbered cell that intersects with the filter, and iii) the perimeter of the filter would be placed such that the minimum number of cells intersect with the filter (Figure 1).

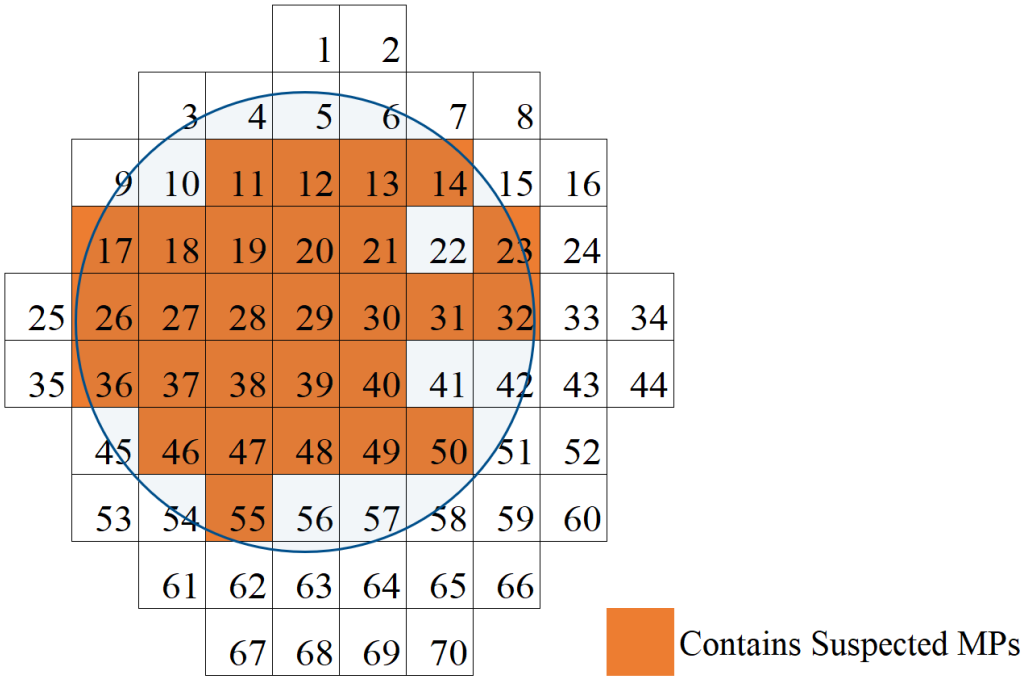


Figure 1: Filter #1 is located following the three rules of i) staying within the numbered cells, ii) incorporating all filters with suspected MPs, and iii) minimizing number of intersecting cells.

If the cell was outside the estimated filter location, a value of zero was given for the relative area of that cell, whereas if the grid cell was inside the estimated filter location, a value of one was assigned. If a cell was both inside and outside the estimated filter perimeter, the proportion of the grid cell inside the filter was approximated using a triangle or rectangle from the vertices of the cell or points of intersection of the perimeter on the cell that best approximated the area of the filter that fell within the grid (Figure 2). The centre cell of the estimated filter location was also recorded. These filters were then visualized in RStudio (RStudio Team, 2018) using software packages *tidyverse* (Wickham, 2019), *readxl* (Wickham et al., 2019), *Rcolorbrewer* (Neuwirth, 2014) and *plot.matrix* (Klinke and Chevalier, 2020).

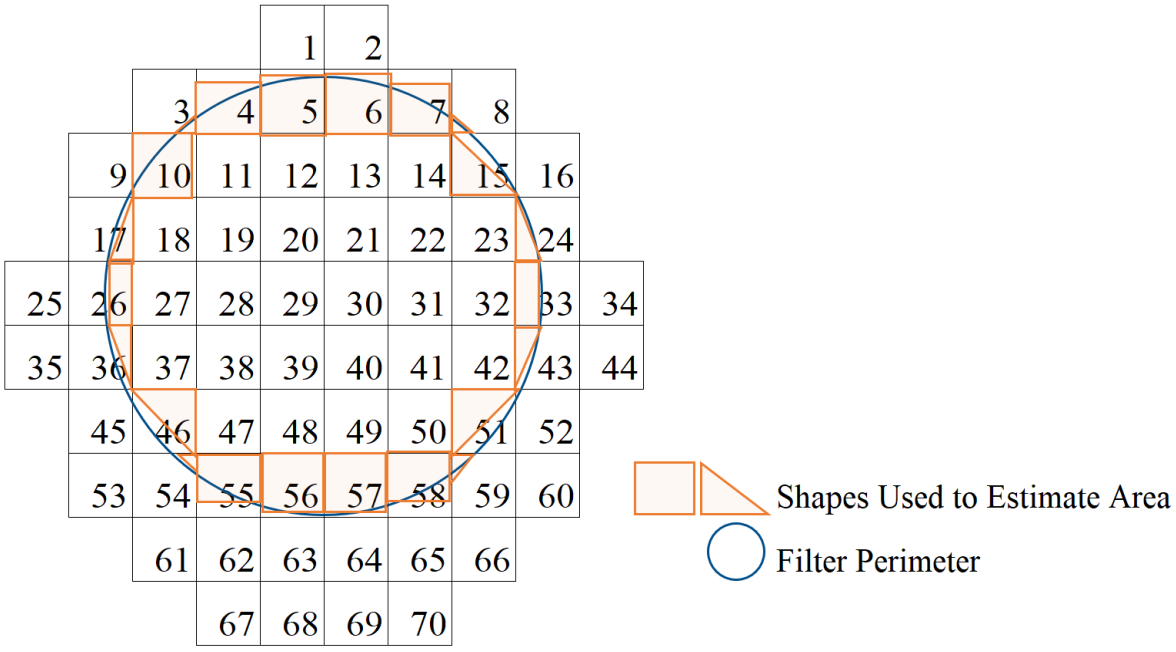


Figure 2: An example of the rectangles and triangles used to estimate the area, whose vertices either touch the intersection of the circle's perimeter or the vertices of the cell, in cells that lie both within and outside of the filter area.

Subsampling Model

In order to predict how much of the filter is required in order to obtain an accurate estimation of the sample population, a bootstrapping methodology was adopted using a similar method to that reported by Jaworski et al. (2019) for statistically justifying sampling areas for mosquito counts. Sampling data were analyzed in RStudio (RStudio Team, 2018) and R (R Core Team, 2013) using additional software packages of *data.table* (Dowle et al., 2020), *rlist* (Ren, 2016), *spatstat* (Baddeley, Turner and Rubak, 2020).

Area-based random sampling methods were used to subsample the suspected MPs. Three sampling methods were used based on similar approaches in the literature (Figure 3). The first approach employed random discontiguous sampling. Grid cells were selected at random within the filter without replacement until a random value between one and all of the grid cells intersecting with

the filter were selected. The second approach employed random contiguous sampling, where one random grid cell was selected, from which adjacent cells were randomly populated cells until every adjacent cell from the initial random point were populated. Then the next concentric ring of cells was randomly populated. This process was repeated until a random value between one and all grid cells of the filter were selected. A final sampling strategy incorporated a two equal cut-out approach, similar to the random contiguous approach except rather than one contiguous shape, two identical contiguous shapes were formed from random cells equidistant from the centre cell in opposite directions.

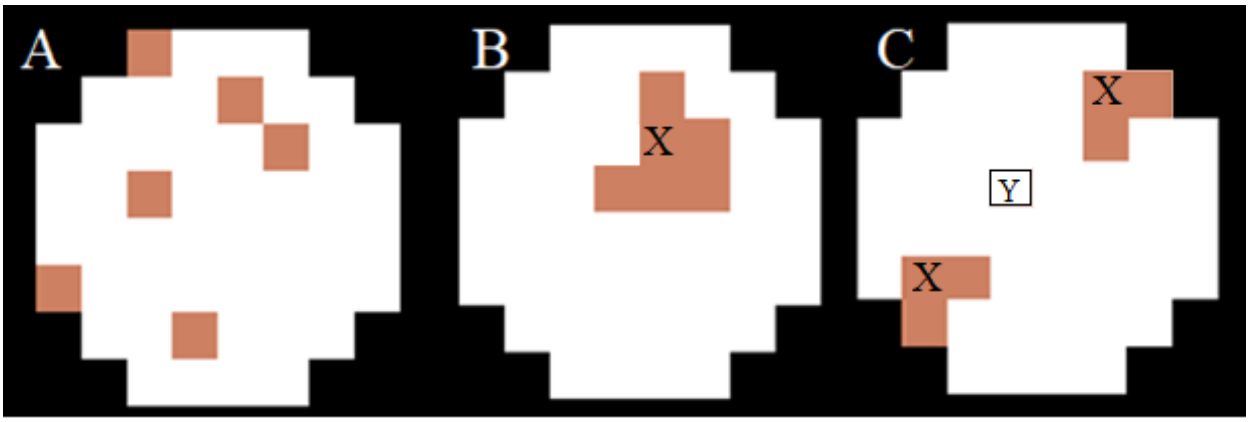


Figure 3: Three different cell subsampling methods: (A) random discontiguous sampling, (B) one contiguous area sampling and (C) two contiguous areas equidistant and opposite directions away from the centre. The first random point is from which concentric circles are circled around shown by X and the centre grid from which the second location is located equidistant apart and opposite in direction is shown by Y. In all three examples, six cells are selected and ~16% of the filter has been sampled.

Individual filters were examined using each sampling strategy 5,000 times without replacement, meaning no two samples per sampling strategy per filter were identical. Estimated total MP count was calculated using Equation 1, where N is the estimated total MPs, n is the total number of MPs in the subsample, A is the total area of the filter, and a is the area of the sample.

$$1) \quad \mathrm{N}_{est} = n \cdot \frac{A}{a}$$

Each estimate was given a consistency score that refers to how far off the estimate is from the true population. Consistency was calculated by dividing the total estimated suspected MP count by the actual total suspected MP population derived from counting the entire filter multiplied by 100. From there, the percentage of total area that was sampled was rounded into intervals of 1%. The 95% confidence interval for the consistency value was determined for each sample size interval of one percent for each filter for each sampling strategy. To answer the question of how much of the filter area is needed to obtain an adequate estimation of the MP population, a suitable subsample filter area was defined as a sample size in which the bulk of the filter estimates lies within a decided upon range around the MP population for that sampling area and every sampling area with greater area than it. Specifically, a suitable range is where a 95% confidence interval exists such that the lower and upper confidence intervals for all filters lie within that predetermined suitable percentage away from true suspected MP population. To examine sensitivity of this approach, the suitable subsampling areas of the filter needed for a range within ±5, ±10 and ±15% of the true population 19 times out of 20 were examined.

To test whether there was non-random behaviour within the filters, a chi-squared test of complete spatial randomness using quadrat counts to fit an inhomogeneous Poisson model was employed. Quadrats of four grid cells were selected such that they exhibited as much detail as possible while also having high enough predicted values to increase the power of the study.

Results

Significant clustering ($p < 0.05$) was observed on every filter except Filters #1 and #11 within the cell quadrats (Appendix 1.1) (Figure 4). Cells containing the highest and lowest suspected MP counts were not typically observed at a particular location relative to the centre. Instead, relatively

small cell regions, located in no particular part of the filter, often held high suspected MP counts, such that as high as 22% of the MPs was observed to be located solely within one filter grid (Figure 4.8). This indicates that the suspected MPs evaluated were not randomly distributed. Evidence of non-randomness indicates that estimated MP count is not independent of subsample area. Furthermore, this indicates that taking 10% of MPs for subsequent analysis and analyzing 10% of the filter are not equivalent processes. Instead, analyzing a portion of the filter and equating it to the percentage of MPs present is only a best estimate, not an underlying characteristic of MP enumeration.

Typically, confidence upper and lower intervals approach a consistency of 100% accuracy as a greater portion of the filter area is evaluated. Specific trends were associated with discontiguous sampling, one area contiguous sampling, and two area contiguous area sampling equidistant apart (Figure 5). When comparing discontinuous sampling to both one and two-area contiguous sampling, acceptable results (±15%) were achieved with a third less of the filter area counted (22% vs 62%) (Figure 6).

For all filters, sampling <10% of the filter resulted in widely varying estimates. When sampling as low as 10% of the filter area, typically consistency values for one contiguous area subsampling (95% CI, 54-146%) fell within ±50%, two area contiguous sampling (95% CI, 63-140%) fell within ±40%, and discontiguous sampling (95% CI, 76-119%) fell within ±20% of the true MP total filter count. From zero to 100% of the filter, discontiguous sampling was consistently more accurate (a narrower confidence interval) when compared to contiguous sampling (Figure 5). Discontinguous sampling of just ≥20% of the population is likely the most efficient approach as greater areas represents a point of diminishing return when considering the required analytical time. Each filter was typically subdivided into a minimum of 37 cells, suggesting that a

discontiguous random selection of eight cells (~22% of the filter) may provide an accurate estimate within ±15% of the entire filter anticipated MP count. However, a minimum of randomly discontiguous sampling of 15 cells (~40%) could be sampled to improve accuracy to within ±10% of the true population.

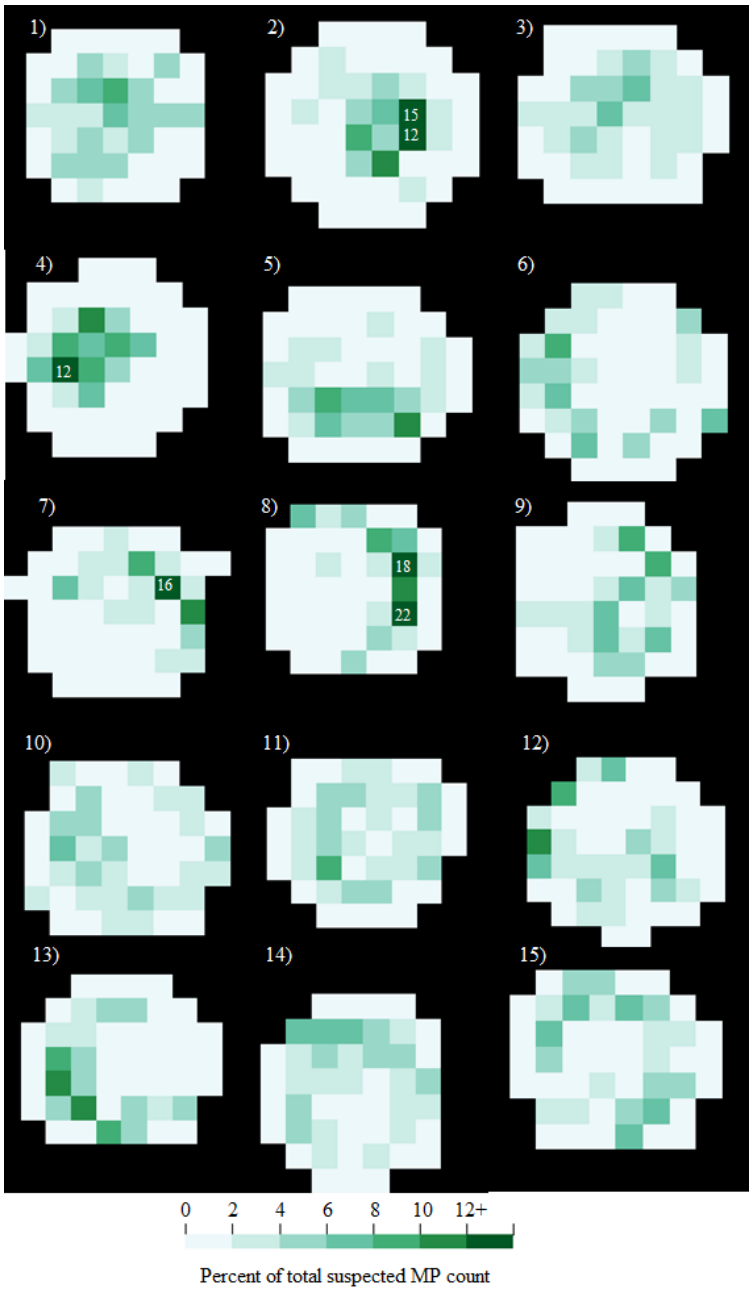


Figure 4: Percent of total suspected MPs for each grid for each completely enumerated filter. Since the suspected MP cell counts are somewhat right skewed, the total MP count in the maximum category is labelled for context.

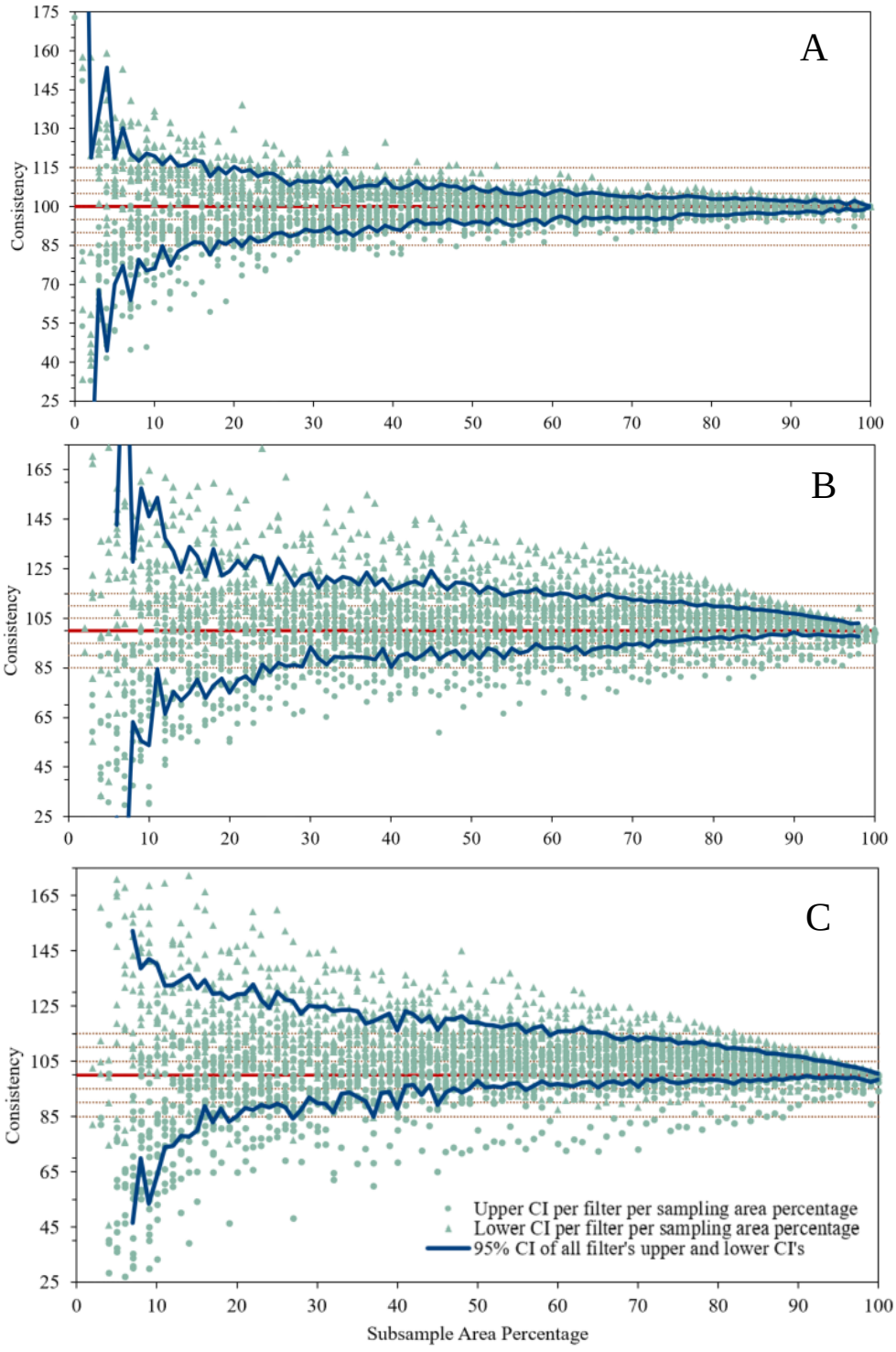
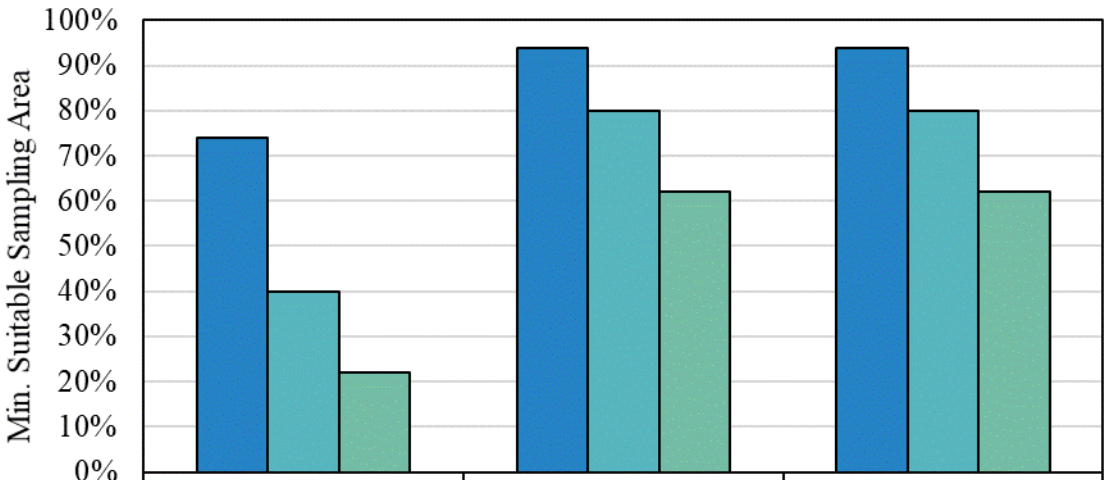


Figure 5: Upper and lower confidence intervals of consistency used in: A) random discontiguous sampling of each filter, B) random contiguous one area sampling, and C) random contiguous one area sampling equidistant and opposite direction from the centre. Values that fell <25% and >175% near a sampling area of 0% were removed to improve visibility.



	Discontiguous	One Contiguous Area	Two Contiguous Areas Equidistant Apart
±5%	74%	94%	94%
±10%	40%	80%	80%
±15%	22%	62%	62%

Figure 6: Minimum area needed to confidently sample a filter in order to be within three possible desired ranges of the true population 19 times out of 20.

When considering only one contiguous sampled region, confidence intervals were larger than those estimated from discontiguous sampling and thus required a much larger area to achieve the same accuracy obtained through discontiguous sampling. To achieve an estimated ±15% of the true population, >60% of the filter area required sampling. The imperfect accuracy observed in filter sampling can be seen as a result of the spatial autocorrelation indicated by significantly non-random distribution of the adjacent four cell quadrats. Cells with high suspected MP counts are likely to be locally surrounded by cells that also have high MP counts. Contiguous sampling amplifies this spatial autocorrelation by sampling nearby cells which overemphasizes clusters and deemphasizes the average values across all regions, resulting in greater inaccuracy.

Contiguous sampling resulted in asymmetrical sampling estimates, where both the higher and lower consistency values are often centered above 100% (Figures 7 and 8). This asymmetry was

observed throughout most of the filters, including Filter #1 in contiguous one area sampling (Figure 7) and, to a lesser extent, in contiguous two area sampling (Figure 8). Filter #1 shows that when sampling 30% or more, estimates for upper and lower confidence intervals exceeded a consistency of 100% (Figure 7). While most filters revealed some level of overestimation, especially ≥50% onwards mark, not all did. In contrast, Filter # 6 along with Filter # 10 show consistent underestimation of the true value when sampling ≥50% in contiguous one area sampling (Figure 7), and to a lesser extent, in contiguous two area sampling (Figure 8). Since random one area contiguous sampling methodology oversamples the centre in comparison to outer cells, an overestimation in the samples results from higher suspected MP counts in the centre. Since the asymmetry shows that there are more filters whose suspected MP count is overrepresented in the centre, one can anticipate that a slightly higher concentration of MPs exists in the centre than on the outer rim, although statistical significance for this was not evaluated.

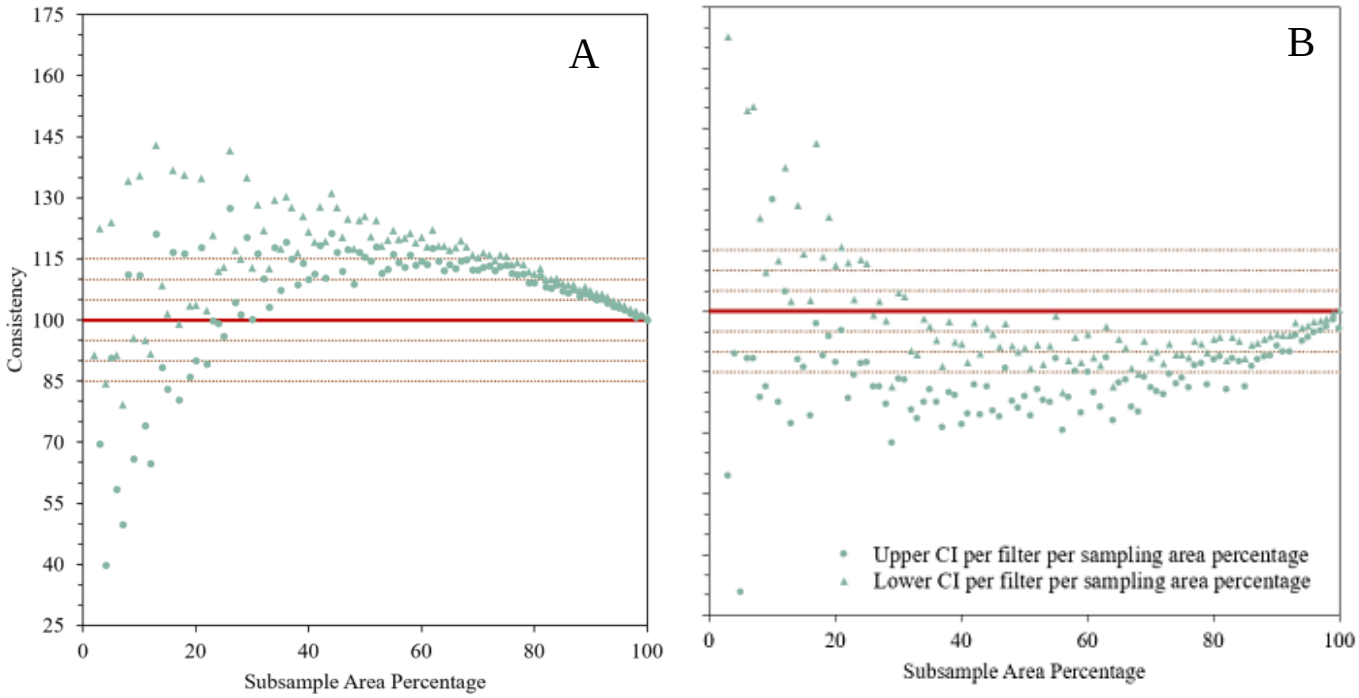


Figure 7: Upper and lower confidence intervals of consistency used in random one area contiguous sampling of A) Filter #1 and B) Filter #2. Values that fell below 25% and above 175% near a sampling area of 0 were removed from the chart for improved visibility.

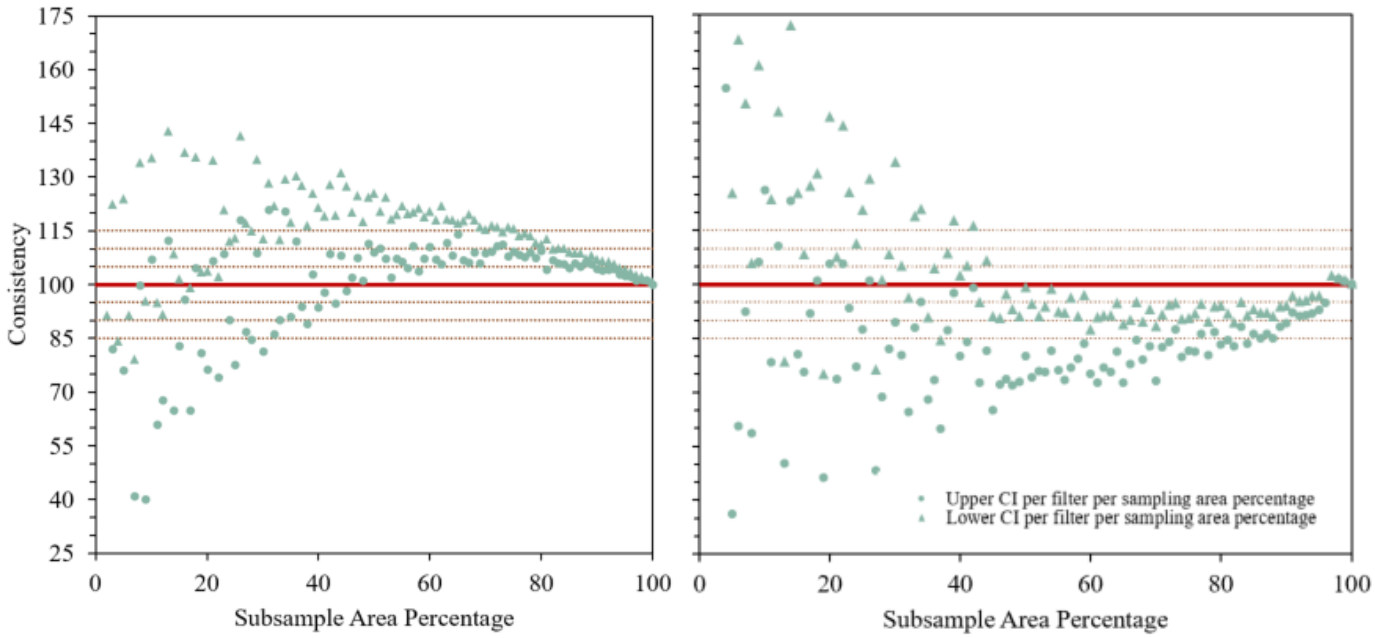


Figure 8: Upper and lower confidence intervals of consistency used in random two area contiguous sampling of A) Filter #1 and B) Filter #2. Values that fell below 25% and above 175% near a sampling area of 0 were removed from the chart for improved visibility.

Sampling in two contiguous areas located equidistant from the centre also caused wider confidence intervals than discontiguous sampling (Figure 6). Overall, sampling in two areas did not improve minimum sampling areas compared to one area contiguous sampling. Clustering must be to such a great extent that having two areas of sampling does not counterbalance the misrepresentation that sampling from concentrated area causes due to cells having similar counts to adjacent cells suspected MP counts.

Having decided that a minimum suitable sample size was defined as having a sample size and every sample size greater than it with an upper and lower confidence interval of consistency within reasonable value (Figure 9), sampling strategies greatly affected how large that sample size needed to be. Out of the 15 filters, the minimum value for suitable sampling area within ±10% was 80% for one and two area discontiguous sampling and half that (40%) for discontiguous sampling (Figure 8). Across all three filters, an increase of sampling estimate accuracy from ± 15% to ± 10% required an additional 18% of the filter area. An increase of sampling accuracy from a consistency range of ± 10% to ± 5% required an added 14% of sampling area for contiguous sampling and an additional 34% of sampling area for discontiguous sampling.

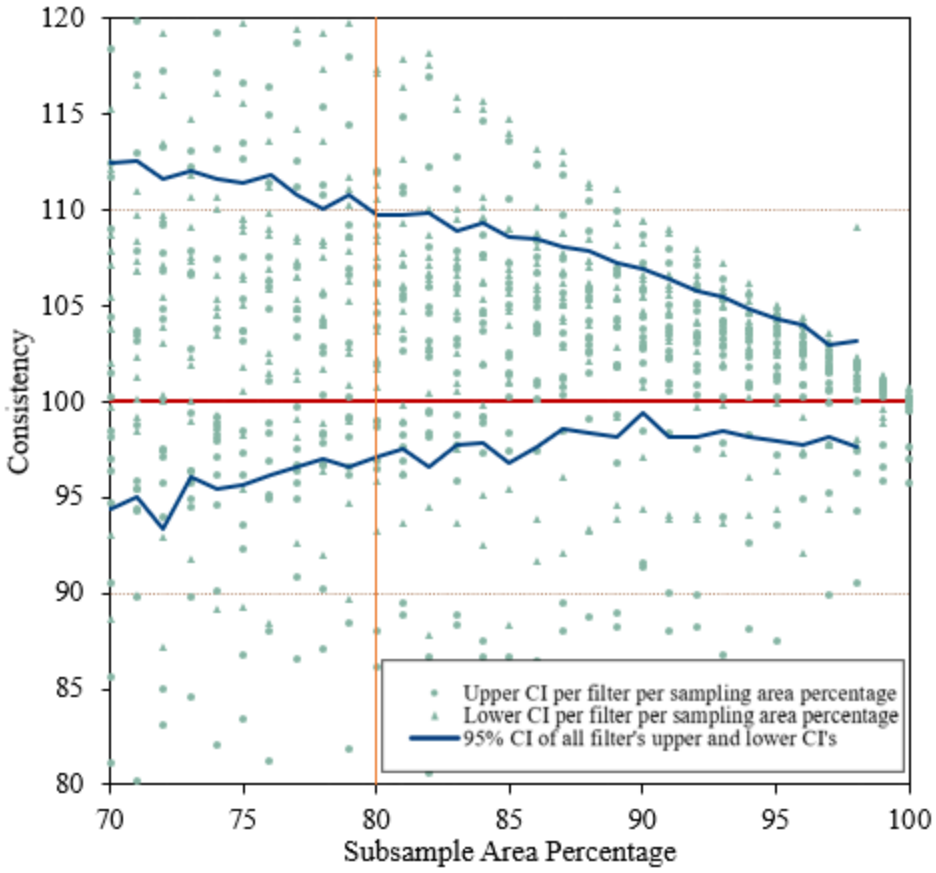


Figure 9: Suitable sample size is given at the point where there is a 95% confidence level that the upper and lower consistency confidence intervals of each filter lie within a predetermined range. For contiguous area sampling, the suitable sample area for an accuracy of 10% is 80% of the filter area. Even though there are still individual confidence intervals that exceed this range, the confidence interval of all of the upper and lower consistencies stays within 10% of true MP estimate when sampling 80% or more of the filter.

Discussion

To address the initial question of how many cells are needed to reliably subsample a MP filter, randomly sampling one in four cells (25% of the filter area) resulted in an adequate estimation within ± 15% of the true suspected MP count 19 times out of 20. When considering published MP literature, consistent issues have been raised regarding the amount of time required to quantify and characterize MPs following filtration (Pivokonský et al., 2020; Imhof et al., 2016). While high-throughput, in-line filtration will ease many concerns regarding total sample size and the QA/QC

protocols needed, the additional MPs that will require analysis as samples increase in volume presents a challenging dilemma: either subsample a higher percentage of the filter area, limiting valuable time that could be used to analyze other samples, or subsample a smaller percentage of the filter area, introducing further uncertainty to the final MP estimate. It is suggested that this difficult trade-off be addressed using statistical rigour via the methods described herein. As such, it is suggested that drinking water researchers examining MPs in a new environment or with an adjusted method perform the following steps to determine how much of the filter area should be subsampled:

1. Determine the smallest desired category of MPs whose concentration requires quantification. Categories maybe include as broad as total MP count to as detailed as counts of MP plastic types, shapes, or sizes.
2. Determine desired MP estimation accuracy, keeping in mind inaccuracy associating with other steps of the methodology.
3. Conduct an initial pilot study whereby the filters surfaces are quantified and characterized entirely. Obtain random samples over a range of different sampling areas to estimate MP counts from those filters.
4. Calculate lower and upper consistency for confidence intervals for estimates.
5. Determine a minimum sampling area whereby the upper and lower 95% confidence intervals between filters is within the desired accuracy and every sampling area greater than it has a confidence level that also lies within the desired accuracy, nineteen times out of twenty.
6. Use the minimum sampling area generated in step 5 as the minimum sampling area.

By employing this method, a greater certainty exists that differences in MP counts in time, space and methodology are true differences rather than uncertainty resulting from sampling. Some of the previous studies show MP counts which exhibit little variation between water treatment plants while only sampling very small portions of the filter (Adib, Mafigholami and Tabeshkia, 2021; Tong et al., 2020). This is concerning as our results suggest that the differences observed in MP counts in studies with small subsample sizes could be explained by the uncertainty introduced by clustering within subsamples.

Spatial evidence of clustering, and observations of density plots reveal that a random distribution for MP on any given filter cannot be assumed. Accurate estimation is dependant on sample size and smaller sample sizes lead directly to increased uncertainty. However, a non-random distribution does not represent a strong correlation between filter location and density. No sampling technique would need to incorporate specific underrepresented areas or diminish overrepresented areas. It is evident that discontiguous sampling represents a superior method when compared to one or even two area contiguous sampling, no matter how much of the filter is being evaluated. It is encouraged that sample strategies incorporate a large number of random cut-outs as a way of minimizing uncertainty, even if this adds time to particle enumeration. Since no strong location dependence exists on MP density, it would not affect uncertainty if the random cells selected at the beginning of sample enumeration be the same cells that are used to subsample each different filter. Two area contiguous sampling provided the exact same results as one area discontiguous sampling when considering suitable sampling area estimations, indicating that a few additional discrete contiguous sampling areas does not improve results and thus is not a replacement for discontiguous sampling. Even though equidistant cut-outs are some of the most common sampling techniques used for counting drinking suspected MPs, this study presents weak

statistical evidence for employing such a sampling strategy unless many cut-outs are used (Pivokonský et al., 2020; Adib, Mafigholami and Tabeshkia, 2021). It is important to recognize that each cell itself can be interpreted as a cut-out. Discontiguous sampling at 22% in the filters in this study is the same as having at least 8 contiguous cut-outs. As such, it is recommended that at least 8 cut-outs are employed to sample MPs on a filter and suggested that the greater sampling area employed, the greater number of randomly selected cut-outs selected.

The consistent overestimation and occasional underestimation in contiguous sampling of suspected MP estimate population were unexpected. The hypothesis for this odd behaviour is that when ≥50% of the filter is enumerated or more, contiguous sampling results in the middle being sampled more than the surrounding areas. The estimates are more heavily sampled toward the centre of filter, and MPs in the centre are more heavily weighted. This explains why Filter #1 with a denser centre overestimates filter count and Filter #6 which has a relatively hollow centre underestimates filter count (Figure 4). Since there were more cells that had a slightly higher MP density in the centre than on the outside, overestimation of MP total count was more likely than underestimating. This overrepresentation of the centre provides further evidence that, when at all possible, discontiguous subsampling should be used for suspected MP enumeration. However, it should be noted that many studies that have employed the cut-out method typically sample less than 50% of the filter area, and are not oversampling the centre (Pittroff et al., 2021; Tong et al., 2020). Our results indicate that sampling in three large central cut-out such as the strategy employed by Pivokonský et al. (2020) may result in some slight overestimation of the filter count and greater inaccuracy in estimation due to clustering of MPs.

While it ultimately depends on the intent of the study, a random discontiguous subsampling area of >22% in order to achieve an accuracy of ± 15% 19 times out of 20 appears to provide an

adequate trade-off between the time needed to subsample and accuracy of the final count when considered treated drinking water. This study not only confirms the one quarter suggestion of filter area analysis suggested by Koelmans et al. (2019), but provides the statistical justification for such a claim. Going forward, this is the recommendation of subsample percent area for studies that may not have the resources to conduct pilot studies.

Next Steps

The crucial next step needed to provide additional statistical rigour when subsampling MPs is to perform the same study with individually identified chemically characterized MPs. If MPs are clustered due to aggregation as opposed to the way in which particles interact with the filter, it is possible that confirmed MPs may be more clustered than suspected MPs (Wang et al., 2021b). The possibility that the true MPs are more clustered and deviate further from a random distribution than suspected MPs would mean that a greater sampling area is needed to achieve the true accuracy within a desired range and is a very legitimate concern that should be evaluated in subsequent research.

A number of other intrinsic parameters that may affect MP distribution, including total MP count, chemical type, size, shape, and origin. Furthermore, methodology variations such as surfactant usage should also be evaluated. When considering in-lab filtration, Cherniak et al. (2020) observed that adding surfactant changes particle behaviour within a sampling vessel, allowing allows more particles to leave the vessel. It would also be beneficial to see how particle distribution on the filter might also be affected by surfactant. While examining how surfactant and total MP count affect sample size accuracy was not the purpose of this study, it should be noted there does not appear to be any abundantly obvious correlation between either total suspected MP count when within the range of total suspected MP examined in this study (45 – 266) nor does there appear to be any

difference between filters with surfactant (Filter #5-15) versus those without (Filter #1-4). It is also possible that source water could potentially influence MP distribution on the filter and merits further research. Since clustering appears to be at the sub-cell layer, understanding MP distribution and how localized clustering behaves would benefit from measuring MP coordinates on the filter rather than the larger cell sizes used in this study.

Conclusion

To answer the question of how much of the filter should be sampled in order to have reasonable confidence in the accuracy of the estimate while also saving the most amount of counting time, this study suggests conducting a pilot study where a framework of repeatedly randomly subsampling from a few filters that are completely counted in order to determine a sample size that fits the parameters of the study. Going forward, subsampling should incorporate many randomly selected cut-outs as that choice of subsampling results in 20-40% less subsampling area needed to obtain the same accuracy of MP estimation when compared to sampling an unbroken subsample area. Our method indicates that a minimum area of a quarter of the filter area is needed to obtain an estimate of the total suspected MP population within ±15% of the true population 19 times out of 20. Additional research is required to determine if this subsampling size is independent of MP characterization, source water, surfactant usage, and manual versus automated enumeration. In general, going forward, studies should attempt to justify the subsampling area chosen as the field attempts to standardize methodology. Until a standardized methodology to assess MP concentration emerges, every step of the MP collection and counting process must be questioned and rigorously justified in order that one method can clearly be accepted as an improvement to another using statistical evidence and convincing substantiation.

References

Adib, D., Mafigholami, R. and Tabeshkia, H., 2021. Identification of Microplastics in Conventional Drinking Water Treatment Plants in Tehran, Iran. *Environmental Science and Pollution Research*, [online] Advance online publication. https://doi.org/10.21203/rs.3.rs-242504/v1.

Adoption of Definition of Microplastics in Drinking Water.[online] Available at: <https://www.waterboards.ca.gov/board_decisions/adopted_orders/resolutions/2020/rs2020_0021.pdf>.

Auta, H.S., Emenike, C.U. and Fauziah, S.H., 2017. Distribution and importance of microplastics in the marine environment: A review of the sources, fate, effects, and potential solutions. *Environment International*, 102, pp.165–176. https://doi.org/10.1016/j.envint.2017.02.013.

Avio, C.G., Gorbi, S., Milan, M., Benedetti, M., Fattorini, D., d'Errico, G., Pauletto, M., Bargelloni, L. and Regoli, F., 2015. Pollutants bioavailability and toxicological risk from microplastics to marine mussels. *Environmental Pollution*, 198, pp.211–222. https://doi.org/10.1016/j.envpol.2014.12.021.

Baddeley, A., Turner, R. and Rubak, E., 2020. *spatstat: Spatial Point Pattern Analysis, Model-Fitting, Simulation, Tests*. [online] Available at: <https://CRAN.R-project.org/package=spatstat> [Accessed 3 Feb. 2021].

Bowley, J., Baker-Austin, C., Porter, A., Hartnell, R. and Lewis, C., 2020. Oceanic Hitchhikers – Assessing Pathogen Risks from Marine Microplastic. *Trends in Microbiology*. [online] https://doi.org/10.1016/j.tim.2020.06.011.

Chen, G., Feng, Q. and Wang, J., 2020. Mini-review of microplastics in the atmosphere and their risks to humans. *Science of The Total Environment*, 703, p.135504. https://doi.org/10.1016/j.scitotenv.2019.135504.

Cherniak, S., Mckie, M., Hemabessiere, L., Yuan, K., Rochman, C. and Andrews, R., 2020. Conventional and Biological Treatment for the Removal of Microplastics from Drinking Water. *Unpublished*.

Cox, K.D., Covernton, G.A., Davies, H.L., Dower, J.F., Juanes, F. and Dudas, S.E., 2019. Human Consumption of Microplastics. *Environmental Science & Technology*, 53(12), pp.7068–7074. https://doi.org/10.1021/acs.est.9b01517.

Deng, Y., Zhang, Y., Lemos, B. and Ren, H., 2017. Tissue accumulation of microplastics in mice and biomarker responses suggest widespread health risks of exposure. *Scientific Reports*, 7(1), p.46687. https://doi.org/10.1038/srep46687.

Dowle, M., Srinivasan, A., Gorecki, J., Chirico, M., Stetsenko, P., Short, T., Lianoglou, S., Antonyan, E., Bonsch, M., Parsonage, H., Ritchie, S., Ren, K., Tan, X., Saporta, R., Seiskari, O., Dong, X., Lang, M., Iwasaki, W., Wenchel, S., Broman, K., Schmidt, T., Arenburg, D., Smith, E., Cocquemas, F., Gomez, M., Chataignon, P., Blaser, N., Selivanov, D., Riabushenko, A., Lee, C., Groves, D., Possenriede, D., Parages, F., Toth, D., Yaramaz-David, M., Perumal, A., Sams, J., Morgan, M., Quinn, M., @javrucebo, @marc-outins, Storey, R., Saraswat, M., Jacob, M., Schubmehl, M., Vaughan, D., Hocking, T., Silvestri, L., Barrett, T., Hester, J., Damico, A., Freundt, S., Simons, D., Andrade, E.S. de, Miller, C., Meldgaard, J.P., Tlapak, V., Ushey, K. and Eddelbuettel, D., 2020. *data.table: Extension of 'data.frame'*. [online] Available at: <https://CRAN.R-project.org/package=data.table> [Accessed 3 Feb. 2021].

Elkhatib, D. and Oyanedel-Craver, V., 2020. A Critical Review of Extraction and Identification Methods of Microplastics in Wastewater and Drinking Water. *Environmental Science & Technology*, 54(12), pp.7037–7049. https://doi.org/10.1021/acs.est.9b06672.

Hartmann, N.B., Rist, S., Bodin, J., Jensen, L.H.S., Nørgaard Schmidt, S., Mayer, P., Meibom, A. and Baun, A., 2017. *Microplastics as Vectors for Environmental Contaminants : Exploring Sorption, Desorption, and Transfer to Biota*. [online] Integrated Environmental Assessment and Management. https://doi.org/10.1002/ieam.1904.

Hwang, J., Choi, D., Han, S., Jung, S.Y., Choi, J. and Hong, J., 2020. Potential toxicity of polystyrene microplastic particles. *Scientific Reports*, 10(1), p.7391. https://doi.org/10.1038/s41598-020-64464-9.

Imhof, H.K., Laforsch, C., Wiesheu, A.C., Schmid, J., Anger, P.M., Niessner, R. and Ivleva, N.P., 2016. Pigments and plastic in limnetic ecosystems: A qualitative and quantitative study on microparticles of different size classes. *Water Research*, 98, pp.64–74. https://doi.org/10.1016/j.watres.2016.03.015.

Jaworski, L., Jansen, S., Pfitzner, W.P., Beck, M., Becker, N., Schmidt-Chanasit, J., Kiel, E. and Lühken, R., 2019. Comparative analysis of subsampling methods for large mosquito samples. *Parasites & Vectors*, 12(1), p.354. https://doi.org/10.1186/s13071-019-3606-5.

Johnson, A.C., Ball, H., Cross, R., Horton, A.A., Jürgens, M.D., Read, D.S., Vollertsen, J. and Svendsen, C., 2020. Identification and Quantification of Microplastics in Potable Water and Their Sources within Water Treatment Works in England and Wales. *Environmental Science & Technology*, 54(19), pp.12326–12334. https://doi.org/10.1021/acs.est.0c03211.

Klinke, S. and Chevalier, F., 2020. *plot.matrix: Visualizes a Matrix as Heatmap*. [online] Available at: <https://CRAN.R-project.org/package=plot.matrix> [Accessed 6 Mar. 2021].

Koelmans, A.A., Mohamed Nor, N.H., Hermsen, E., Kooi, M., Mintenig, S.M. and De France, J., 2019. Microplastics in freshwaters and drinking water: Critical review and assessment of data quality. *Water Research*, 155, pp.410–422. https://doi.org/10.1016/j.watres.2019.02.054.

Li, J., Liu, H. and Paul Chen, J., 2018. Microplastics in freshwater systems: A review on occurrence, environmental effects, and methods for microplastics detection. *Water Research*, 137, pp.362–374. https://doi.org/10.1016/j.watres.2017.12.056.

Lu, L., Wan, Z., Luo, T., Fu, Z. and Jin, Y., 2018. Polystyrene microplastics induce gut microbiota dysbiosis and hepatic lipid metabolism disorder in mice. *The Science of the Total Environment*, 631–632, pp.449–458. https://doi.org/10.1016/j.scitotenv.2018.03.051.

Mintenig, S.M., Löder, M.G.J., Primpke, S. and Gerdts, G., 2019. Low numbers of microplastics detected in drinking water from ground water sources. *Science of The Total Environment*, 648, pp.631–635. https://doi.org/10.1016/j.scitotenv.2018.08.178.

Neuwirth, E., 2014. *RColorBrewer: ColorBrewer Palettes*. [online] Available at: <https://CRAN.R-project.org/package=RColorBrewer> [Accessed 6 Mar. 2021].

Novotna, K., Cermakova, L., Pivokonska, L., Cajthaml, T. and Pivokonsky, M., 2019. Microplastics in drinking water treatment – Current knowledge and research needs. *Science of The Total Environment*, 667, pp.730–740. https://doi.org/10.1016/j.scitotenv.2019.02.431.

Panno, S.V., Kelly, W.R., Scott, J., Zheng, W., McNeish, R.E., Holm, N., Hoellein, T.J. and Baranski, E.L., 2019. Microplastic Contamination in Karst Groundwater Systems. *Groundwater*, 57(2), pp.189–196. https://doi.org/10.1111/gwat.12862.

Pittroff, M., Müller, Y.K., Witzig, C.S., Scheurer, M., Storck, F.R. and Zumbülte, N., 2021. Microplastic analysis in drinking water based on fractionated filtration sampling and Raman microspectroscopy. *Environmental Science and Pollution Research*. [online] https://doi.org/10.1007/s11356-021-12467-y.

Pittura, L., Avio, C.G., Giuliani, M.E., d'Errico, G., Keiter, S.H., Cormier, B., Gorbi, S. and Regoli, F., 2018. Microplastics as Vehicles of Environmental PAHs to Marine Organisms: Combined Chemical and Physical Hazards to the Mediterranean Mussels, Mytilus galloprovincialis. *Frontiers in Marine Science*, [online] 5. https://doi.org/10.3389/fmars.2018.00103.

Pivokonsky, M., Cermakova, L., Novotna, K., Peer, P., Cajthaml, T. and Janda, V., 2018. Occurrence of microplastics in raw and treated drinking water. *Science of The Total Environment*, 643, pp.1644–1651. https://doi.org/10.1016/j.scitotenv.2018.08.102.

Pivokonský, M., Pivokonská, L., Novotná, K., Čermáková, L. and Klimtová, M., 2020. Occurrence and fate of microplastics at two different drinking water treatment plants within a river catchment. *Science of The Total Environment*, 741, p.140236. https://doi.org/10.1016/j.scitotenv.2020.140236.

Qi, R., Jones, D.L., Li, Z., Liu, Q. and Yan, C., 2020. Behavior of microplastics and plastic film residues in the soil environment: A critical review. *Science of The Total Environment*, 703, p.134722. https://doi.org/10.1016/j.scitotenv.2019.134722.

R Core Team, 2013. *R: A language and environment for statistical computing.* [online] Vienna, Austria: R Foundation for Statistical Computing. Available at: <http://www.R-project.org/>.

Ren, K., 2016. *rlist: A Toolbox for Non-Tabular Data Manipulation*. [online] Available at: <https://CRAN.R-project.org/package=rlist> [Accessed 3 Feb. 2021].

Rist, S., Carney Almroth, B., Hartmann, N.B. and Karlsson, T.M., 2018. A critical perspective on early communications concerning human health aspects of microplastics. *Science of The Total Environment*, 626, pp.720–726. https://doi.org/10.1016/j.scitotenv.2018.01.092.

Rivers, M.L., Gwinnett, C. and Woodall, L.C., 2019. Quantification is more than counting: Actions required to accurately quantify and report isolated marine microplastics. *Marine Pollution Bulletin*, 139, pp.100–104. https://doi.org/10.1016/j.marpolbul.2018.12.024.

RStudio Team, 2018. *RStudio: Integrated Development for R.* [online] Boston, MA. Available at: <www.rstudio.com>.

Sarkar, D.J., Das Sarkar, S., Das, B.K., Praharaj, J.K., Mahajan, D.K., Purokait, B., Mohanty, T.R., Mohanty, D., Gogoi, P., Kumar V, S., Behera, B.K., Manna, R.K. and Samanta, S., 2021. Microplastics removal efficiency of drinking water treatment plant with pulse clarifier. *Journal of Hazardous Materials*, 413, p.125347. https://doi.org/10.1016/j.jhazmat.2021.125347.

Shruti, V.C., Pérez-Guevara, F. and Kutralam-Muniasamy, G., 2020. Metro station free drinking water fountain- A potential 'microplastics hotspot' for human consumption. *Environmental Pollution (Barking, Essex: 1987)*, 261, p.114227. https://doi.org/10.1016/j.envpol.2020.114227.

Silva, A.B., Bastos, A.S., Justino, C.I.L., da Costa, J.P., Duarte, A.C. and Rocha-Santos, T.A.P., 2018. Microplastics in the environment: Challenges in analytical chemistry - A review. *Analytica Chimica Acta*, 1017, pp.1–19. https://doi.org/10.1016/j.aca.2018.02.043.

Strand, J., Feld, L., Murphy, F., Mackevica, A. and Hartmann, N.B., 2018. *Analysis of microplastic particles in Danish drinking water*. [Scientific Report from DCE – Danish Centre for Environment and Energy] Aarhus University.p.39. Available at: <https://backend.orbit.dtu.dk/ws/portalfiles/portal/187620768/SR291.pdf>.

Tong, H., Jiang, Q., Hu, X. and Zhong, X., 2020. Occurrence and identification of microplastics in tap water from China. *Chemosphere*, 252, p.126493. https://doi.org/10.1016/j.chemosphere.2020.126493.

Van Cauwenberghe, L. and Janssen, C.R., 2014. Microplastics in bivalves cultured for human consumption. *Environmental Pollution*, 193, pp.65–70. https://doi.org/10.1016/j.envpol.2014.06.010.

Wang, G., Lu, J., Li, W., Ning, J., Zhou, L., Tong, Y., Liu, Z., Zhou, H. and Xiayihazi, N., 2021a. Seasonal variation and risk assessment of microplastics in surface water of the Manas River Basin, China. *Ecotoxicology and Environmental Safety*, 208, p.111477. https://doi.org/10.1016/j.ecoenv.2020.111477.

Wang, X., Bolan, N., Tsang, D.C.W., Sarkar, B., Bradney, L. and Li, Y., 2021b. A review of microplastics aggregation in aquatic environment: Influence factors, analytical methods, and environmental implications. *Journal of Hazardous Materials*, 402, p.123496. https://doi.org/10.1016/j.jhazmat.2020.123496.

Wang, Z., Lin, T. and Chen, W., 2020. Occurrence and removal of microplastics in an advanced drinking water treatment plant (ADWTP). *Science of The Total Environment*, 700, p.134520. https://doi.org/10.1016/j.scitotenv.2019.134520.

Wesch, C., Elert, A.M., Wörner, M., Braun, U., Klein, R. and Paulus, M., 2017. Assuring quality in microplastic monitoring: About the value of clean-air devices as essentials for verified data. *Scientific Reports*, 7(1), pp.1–8. https://doi.org/10.1038/s41598-017-05838-4.

WHO, 2019. *Microplastics in drinking-water*. [online] World Health Organization. Available at: <http://www.who.int/water_sanitation_health/publications/microplastics-in-drinking-water/en/> [Accessed 23 Oct. 2020].

Wickham, H., 2019. *tidyverse*. [online] Available at: <https://CRAN.R-project.org/package=tidyverse> [Accessed 23 Mar. 2020].

Wickham, H., Bryan, J., Komarov, V., Leitenne, C., Colbert, B., Hoerl, D. and Miller, E., 2019. *readxl: Read Excel Files*. [online] Available at: <https://CRAN.R-project.org/package=readxl> [Accessed 27 Mar. 2020].

Yuan, C., 2021. *Develop a Sampling and Extraction Method for Microplastics in Drinking Water to Investigate the Impact of Ultrafiltration on Microplastic Removal or Addition*. Civil Engineering, University of Toronto.

Appendix

Table A.1: Statistical parameters for clustering of suspected MPs on each filter

Filter Number	Degrees of Freedom	Significance of Clustering (p value)
1	20	0.06893
2	15	< 0.00001*
3	20	0.00026*
4	15	< 0.00001*
5	15	< 0.00001*
6	15	< 0.00001*
7	16	< 0.00001*
8	15	0.00002*
9	15	< 0.00001*
10	16	< 0.00001*
11	20	0.08326
12	15	< 0.00001*
13	20	< 0.00001*
14	15	< 0.00001*
15	15	0.00044*
Notes: *p-value < 0.05		

Table A.2: Total suspected MPs counts on Filter

Filter Number	Total suspected MPs
1	74
2	85
3	266
4	83
5	177
6	273
7	177
8	45
9	104
10	123
11	204
12	79
13	196
14	223
15	190