Ecotoxicology and Environmental Safety 275 (2024) 116243

Contents lists available at ScienceDirect

Ecotoxicology and Environmental Safety

journal homepage: www.elsevier.com/locate/ecoenv

How many microplastics do you need to (sub)sample?

Win Cowger [a,c,*], Laura A.T. Markley [b], Shelly Moore [a], Andrew B. Gray [c], Kshitij Upadhyay [d], Albert A. Koelmans [e]

[a] Moore Institute for Plastic Pollution Research, 120 N Marina Drive Long Beach, CA 90803, USA
[b] Syracuse University, Syracuse, USA
[c] University of California, Riverside, USA
[d] National Institute of Technology, Raipur, India
[e] Wageningen University, Aquatic Ecology and Water Quality Management Group, PO Box 47, Wageningen 6700 AA, the Netherlands

ARTICLE INFO

Edited by Dr. R. Pereira

Keywords:
Microplastics
Detection limits
Uncertainty
Error
Statistics

ABSTRACT

Analysis of microplastics in the environment requires polymer characterization as a confirmation step for suspected microplastic particles found in a sample. Material characterization is costly and can take a long time per particle. When microplastic particle counts are high, many researchers cannot characterize every particle in their sample due to time or monetary constraints. Moreover, characterizing every particle in samples with high plastic particle counts is unnecessary for describing the sample properties. We propose an a priori approach to determine the number of suspected microplastic particles in a sample that should be randomly subsampled for characterization to accurately assess the polymer distribution in the environmental sample. The proposed equation is well-founded in statistics literature and was validated using published microplastic data and simulations for typical microplastic subsampling routines. We report values from the whole equation but also derive a simple way to calculate the necessary particle count for samples or subsamples by taking the error to the power of negative two. Assuming an error of 0.05 (5 %) with a confidence interval of 95 %, an unknown expected proportion, and a sample with many particles (> 100k), the minimum number of particles in a subsample should be 386 particles to accurately characterize the polymer distribution of the sample, given the particles are randomly characterized from the full population of suspected particles. Extending this equation to simultaneously estimate polymer, color, size, and morphology distributions reveals more particles (620) would be needed in the subsample to achieve the same high absolute error threshold for all properties. The above proposal for minimum subsample size also applies to the minimum count that should be present in samples to accurately characterize particle type presence and diversity in a given environmental compartment.

1. Introduction

Concentrations of microplastics in environmental samples can vary by many orders of magnitude (Pabortsava and Lampitt, 2020). This makes collecting and analyzing a representative sample challenging (Sutton, 2019). Extremely low (< 10) (Miller et al., 2017) and high (> 1000) (Qian et al., 2024) particle counts are common. Low particle counts may be below detection thresholds (Nel et al., 2021; Weber and Kerpen, 2023) and high particle counts may be too numerous to analyze. Low particle counts will less accurately characterize the distribution of plastic properties within the environmental compartment under study (De Frond et al., 2023). Manual picking of individual particles from samples and manual spectroscopic analysis are among the most common practices in the field (Cowger et al., 2020; Primpke et al., 2020a, 2020b). Microplastic particles assessed are often on the larger size of the microplastic particle sizes (5–0.5 mm), which have only recently been automated (Cowger et al., 2024), and automated analysis techniques for smaller particles require expensive equipment. Plastic pollutants are diverse (Rochman et al., 2019), and researchers commonly characterize every particle by polymer, size, morphology, and color (Roscher et al., 2021). Identifying all particles individually in samples is time-intensive and expensive (at the Moore Institute ~1 week and ~1500$ per sample), and can be statistically unnecessary to meet most study objectives if particle counts are too numerous (Koelmans et al., 2019; Mintenig et al., 2020; Primpke et al., 2020a, 2020b).

Subsample size recommendations have been made as 5–10 % of

* Corresponding author at: Moore Institute for Plastic Pollution Research, 120 N Marina Drive Long Beach, CA 90803, USA.
E-mail address: wincowger@gmail.com (W. Cowger).

https://doi.org/10.1016/j.ecoenv.2024.116243
Received 19 January 2024; Received in revised form 8 March 2024; Accepted 18 March 2024
Available online 23 March 2024

particles (European Commission., 2013), up to 20 particles per morphology (Werbowski et al., 2021), 7000 particles total (Brandt et al., 2021), or 12–50 % of a sample filter (Brandt et al., 2021; El Khatib et al., 2023; Mintenig et al., 2020). New studies have begun to calculate the uncertainty associated with subsampling procedures, e.g. Mintenig et al. (2020) reported a coefficient of variance from 0.02 to 0.53 for polymer proportions when analyzing 75 % of a filter and De Frond et al. (2023) calculated that 75 particles could be used to characterize polymer proportions with a summed error of 20 % based on resampling from fully characterized samples. Although percent-based thresholds are proposed in many studies, Brandt et al. (2021) demonstrated that percent-based thresholds had higher error across samples than particle count based subsampling, especially for low count samples, suggesting that total subsample particle count was the driving force behind statistical error. Additionally, from first principles, we can deduce that a sample with, for example hundreds of millions of particles, wouldn't need anywhere near the percent based thresholds reported above for adequate characterization of the major particle types. Therefore we focus this work on the absolute particle counts necessary for adequate subsamples.

Although previous studies have come to similar conclusions about necessary particle counts, statistical underpinnings have been simulation-driven (Brandt et al., 2021), and there does not exist an a priori equation for calculating the appropriate number of particles in samples and subsamples for one's study goals or calculating uncertainty due to particle counts. Furthermore, all previous studies have assumed univariate inquiry (e.g., all a researcher wants to know is polymer type). In contrast, microplastic researchers often want to know the proportions of 4 or more particle properties. This type of inquiry requires a correction for multiple comparisons. A priori analysis happens before data is collected or in the absence of a complete dataset and is often used in power analysis for determining the required sample size to meet study uncertainty goals. An a priori approach with deep statistical underpinning (Brandt et al., 2021) would allow researchers to budget for sample processing costs and personnel time, streamline method development priorities, and establish high throughput techniques.

This study aimed to develop an a priori procedure for quantifying the minimum particle count required to characterize microplastic properties in environmental samples accurately. To this end, we validate well-founded statistical equations for estimating mean proportions using real environmental microplastic data and simulations. We use these equations to investigate the parameter space for a microplastic sample and subsample counts. Based on these equations, we propose an overarching strategy for collecting representative environmental samples and subsampling from collected samples. Lastly, we propose the next steps for the field to continue improving representative sampling techniques.

2. Methodology

2.1. Mathematical foundations and extension

Establishing estimates of population proportions for a subsample is not new and can be effectively applied to the challenges present in microplastic research. Statisticians have used such equations for decades (Daniel and Cross, 2018). In biostatistics, these equations are primarily used to determine appropriate sample sizes for human health surveys. Instead of asking humans about their populations, we will ask microplastic particles about their populations. The essential foundation for estimating sample size begins with the case where the population the researcher is sampling from is very large (Eq. (1)).

$$SSu = round\left(\frac{\alpha^2 * Pe * (1 - Pe)}{E^2}\right) \tag{1}$$

SSu is the uncorrected sample size estimated by the equation. The critical value (α) is the two-tailed critical value for the confidence interval (typically 95 % with a z-score critical value of 1.96) on the error (E). The error is the maximum difference between the true probability of the particle type and the subsampled probability. A generally acceptable value for this is less than or equal to 0.05. The implications of using this value are that the researcher is less likely to observe particle types if they are in lower abundance than 5 % of the sample, and the error is roughly the confidence interval if the measured value is equal to the expected probability (*Pe*) so differences between proportions of less than 10 % may not be significantly differentiated at $p < 0.05$. The expected probability of the property of interest is typically set to 0.5 (50 %) when there is no prior understanding of the likely probability, which maximizes the count needed to get the desired error. Most of the time, researchers should use 0.5 (or a likely expected probability closest to it from all types in the property of interest) since highly variable plastic type proportions are reported in the literature, for example, by color (Martí et al., 2020), morphology (Guo et al., 2021), and polymer (Erni-Cassola et al., 2019). The result is rounded to the nearest integer since fractions of particles cannot be assessed. In the case where the population is small, we cannot ignore the effects of the population size on the outcome and must additionally use Eq. (2).

$$SSc = round\left(\frac{Pop * SSu}{SSu + Pop - 1}\right) \tag{2}$$

The uncorrected population size is input to the equation with the number of plastic particles expected to be in the sample (*Pop*). Lastly, the result is rounded to the nearest whole number to calculate the corrected sample size (*SSc*).

Extending these equations further to another use case for microplastics research is characterizing multiple microplastic properties simultaneously. For example, a researcher may want to randomly (sub) sample particles to characterize color, size, polymer, and morphology simultaneously. This type of analysis necessitates a correction factor for multiple hypothesis tests, and we decided to use the Šidák method (Vickerstaff et al., 2019) to calculate an adjusted critical value (αj) (Eq. (3)).

$$\alpha j = zscore\left(\frac{1 - CI^{\frac{1}{g}}}{2}\right) \tag{3}$$

The confidence interval (*CI*) is taken to the power of 1 divided by the number of groups (*g*). The rest of the equation proceeds normally for calculating the critical value with a two-tailed z-score. Essentially, the researcher needs a higher confidence than one test because of family-wise errors that can occur when multiple independent tests are performed, and all must reach the same confidence. Using this extension to the equation is highly conservative because if groups are dependent (e.g., the percent of the color blue is correlated with the percent of polyethylene) to a high degree, then the multiple tests effect goes down and approaches the traditional z-score calculation. To correct for correlation between variables, we can adjust the group value (*gc*) by the level of correlation using the Dubey/Armitage-Parmar method (Vickerstaff et al., 2019) (Eq. 4).

$$gc = g^{(1-c)} \tag{4}$$

The number of groups is exponentiated by 1 minus the mean correlation between groups (*c*). This formulation is intuitive because when variables are perfectly correlated, the number of groups becomes set to 1 and when variables are perfectly uncorrelated, the number of groups is unchanged. In most cases, a user will not know what the mean correlation between groups is until after the data has been collected, in a priori cases (the focus of this study) we recommend taking a conservative approach (the approach taken for this study) and setting the correlation to zero or not using this adjustment.

Considering that these equations are not simple to set up, we set out to summarize the equation for rough calculations in the field. If we fit the proposed values to Eq. (1), we can derive a highly simplified version of the sample size equation.

$$SSsu = E^{-2} \tag{5}$$

The simplified uncorrected sample size (SSsu) can be estimated using the error alone to the negative power of 2. Similarly, this equation could calculate the error with a given sample size by taking the sample size to the negative one-half power.

2.2. Environmental validation and simulation parameters

We validated Eqs. (1)–(3) using simulation and environmental measurements. Environmental measurements came from the study Nava et al. (2023) of 117 samples of microplastics on the surface of lakes worldwide. From each sample, we randomly selected 10–200 (in increments of 10) particles and calculated the error for each property of the subsample. We took the 95 % quantile of the errors for all properties in the sample and then used the maximum of the high absolute errors out of the color or morphology values as the input for error to the equations. We compared the actual subsample count with the equation-derived count by including the default proportion of 0.5, 95 % confidence interval, the total microplastic count in the samples, and 2 groups.

Next, we simulated a realistic range of microplastic samples and their polymer distribution, subsampled from them randomly, and derived the high absolute error for the polymer distribution of the subsample from the sample (Fig. 1). We then input the simulation values for error, confidence intervals, groups, and no expected correlation (0) into Eqs. (1)–(4) to predict the sample size used in the simulation. Lastly, we compared the predicted sample size from the equation to the simulated sample size in a 1:1 plot.

Imagine the diversity of sample polymer distributions that could be in the environment. Typically, 2–10 polymer types are found in a sample, and often the polymer abundances are skewed, with 2 or 3 of the polymer types being in the highest abundance (Kooi et al., 2021; Kooi and Koelmans, 2019; Roscher et al., 2022). Samples can have zero to thousands of suspected microplastic particles (Moore et al., 2011). Potential subsamples can theoretically range from a single particle to the entire contents of the sample. These considerations are similar for microplastic polymer, size, color, and morphology quantification. We used these facts to set up our simulation scenario (Fig. 1).

Simulations were conducted in R (RStudio Team, 2020) with ggplot (Wickham, 2016), dplyr (Wickham et al., 2020), tidyr (Wickham and Girlich, 2022), and data.table (Dowle and Srinivasan, 2020) packages. We used a uniform distribution to choose the number of polymers in the simulated sample from 2 to 10. We generated Poisson distributed percentages for the suite of polymer types with a Poisson sampler with lambda = 1 and added 1 to all Poisson sampled percentages. Then, we transformed the values using the softmax function (divided all values by the sum of the values). The sample particles were generated using the probabilities of the polymer types as weights in the sample function and varying the sample count from 10 to 1,000,000 by order of magnitude steps. Subsamples within the simulations ranged from 1 particle to the total number of particles in the sample by order of magnitude steps. Subsampling was conducted using a uniform sampling distribution (De Frond et al., 2023). In practice, the uniform distribution randomly samples from all particles, so using the output from this simulation assumes that methodology is followed. The percentages for each polymer type in the sample were compared to those in the subsample, and the high absolute error was derived by taking the absolute difference between the paired polymer type percentages and estimating their 95th quantile. This process was repeated 100 times for each pair of sample counts and subsample counts. The 95th quantile of the 100 high absolute errors for each pair was extracted (error). This simulation was extended

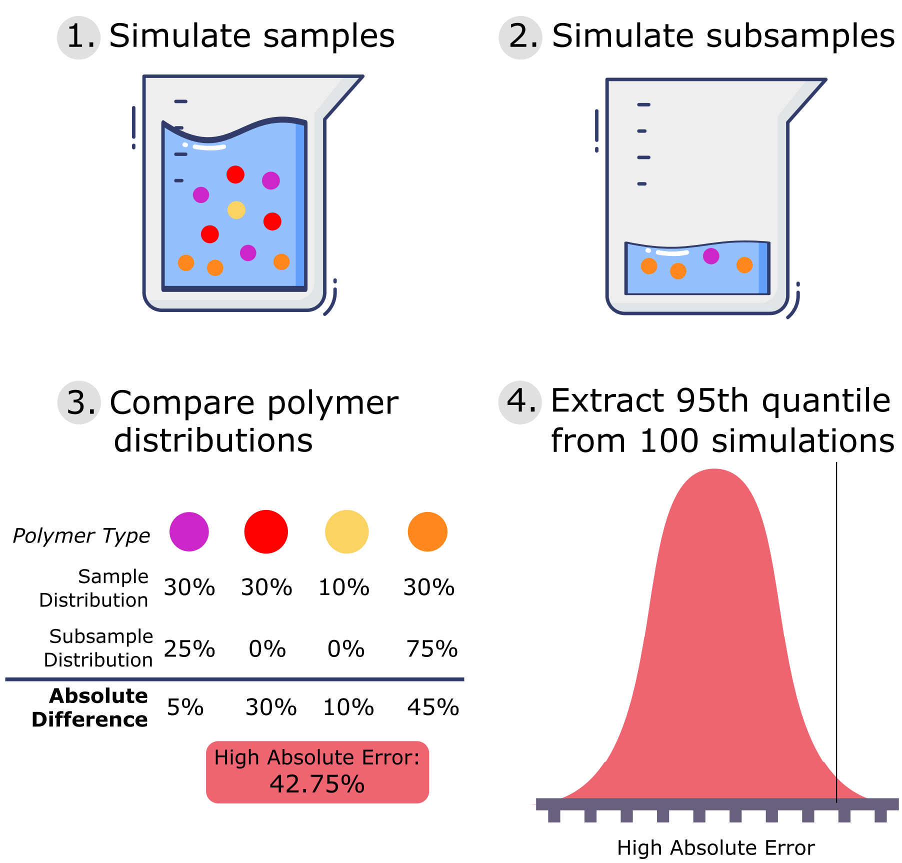

Fig. 1. The simulation setup. 1) represents a theoretical sample with a population of particle types, 2) represents a subsample from that sample, 3) represents a single simulation result and estimation of high absolute error for the particle property based on the 95 % quantile from the absolute differences, 4) shows a probability distribution for all simulations and the 95 % quantile was used as the error estimate.

to include a subsample that could account for particle size, morphology, and color by repeating the same simulation above three additional times for each particle subsample pair and quantifying the error.

3. Results and Discussion

3.1. Validation results

We tested the utility of the Eqs. (1)–(4) for estimating common composition domains of microplastic samples using simulations (Fig. 1). The measured/simulated and calculated estimates for sample size had a nearly a 1:1 correspondence for the estimates for single particle properties (e.g., polymers) (Fig. 2). The lack of a perfect fit was because the simulations are random and biased toward the microplastic scenarios we modeled. The equation output was always conservative leading to slightly larger sample sizes than required in most cases (42 % higher on average). The ratio between measured and calculated sample counts ranged from 0.37 – 1.08 with a mean of 0.70 and a median of 0.71. Only 1 of the 62 values ended up having an underestimate (by 8 %) from the mathematical equation which is within the underestimates expected based on 95 % confidence intervals. The estimates for multiple groups (red dots and green dots) performed similarly to the single group which is evidence in favor of using the adaptations to Eq. (3).

3.2. Subsampling particle properties

We explored the parameter space using Eqs. 1–3, altering error from 10^{-5} to 10^{-1} by order of magnitude steps, sample size from 10^1 to 10^{10}, and 1 or 4 groups (Fig. 3). The parameter space shows that, in general, sample size is less important than high absolute error in estimating the subsample size needed, except for when the sample size approaches the subsample size when the error trends toward negative infinity. This plot could be used to estimate one parameter when others are known. It is important to note that error also relates to the confidence interval of the proportions in the sample. For example, if the researcher finds 0.5 proportion of polyethylene in their sample at 0.05 error, then the confidence interval for the proportion ranges from 0.45 to 0.55. Similarly, particle property proportion differences of less than 10 % might be indistinguishable due to the overlap in their potential estimates. The confidence interval technically decreases for smaller proportions, but this rough guide should be helpful. The error metric can also guide the detection limit for low-abundance particle properties. For example, if the error threshold is set at 5 %, the subsampling approach may lose particle properties in abundances lower than 5 %. Extending the equation to 4 groups simultaneously reveals that slightly higher subsample counts would be needed to accurately characterize the proportions of all particle properties given the same error and sample size.

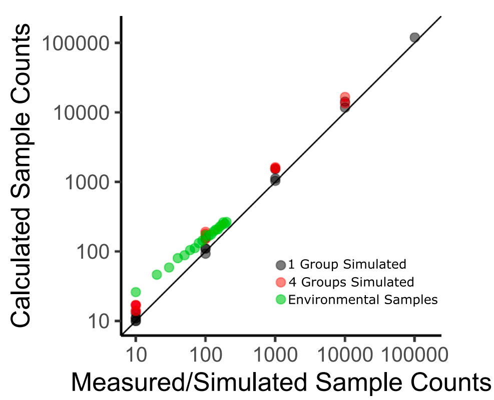

Fig. 2. Validation of the equations. Red dots are the estimated number of particles needed for a sample based on simulation (x axis) and the mathematical formula (y axis) when considering 4 independent groups simultaneously. The black dots are the simulated and mathematical values for single groups. Green dots are for relations derived using 117 environmental samples from Nava et al. (2023) with two groups (color and morphology). The black line through the center is the 1:1 line. Simulations and calculations used 95 % confidence intervals. Error values varied depending on the subsample size compared to the population size and error values from the simulations were used to perform the calculations in the equations to estimate the sample size.

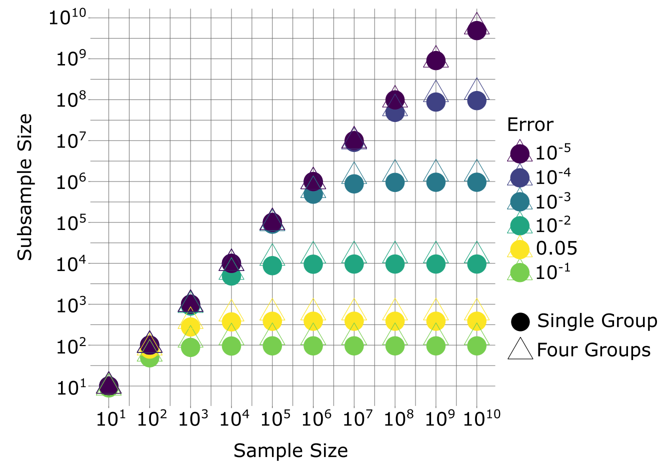

Fig. 3. The total number of plastic particles in the sample (x axis) and number of particles needed to subsample (y axis) given the desired amount of error (colors) for samples where single groups are being analyzed (circles) and where 4 groups are being analyzed simultaneously (triangles). This figure can be used as a quick reference guide for researchers planning to subsample for microplastics. Eqs. (1)–(3) were used to make this figure. The critical value was set to 1.96 and the expected probability was set to 0.5. All other parameters were varied, as shown in the figure.

3.3. How large should samples be to represent their compartment?

All samples of compartments are subsamples. It is typically impractical to sample an entire matrix of interest in a single study due to the size of the compartment and the number of plastic particles in it (Eriksen et al., 2023). For example, if one aims to estimate the polymer distribution of plastic particles in the river flow at a cross-section, a subsample must be made (Cowger et al., 2022, 2021). Therefore, one could call samples subsamples of the compartments. These equations apply no matter the compartment size, a small pond, or the entire earth.

We can use Eqs. (1) and (3) to determine the number of particles that should be in a sample when it is to be representative of a matrix with many microplastic particles > 100,000 (within the sample time and space domain). Assuming that we want to measure particles with 4 properties with an error of 5 % or less, we estimate that samples should contain 620 plastic particles or more (Fig. 4). If we only care about subsampling a single property, we would estimate that we should collect at least 384 plastic particles. The minimum error we hope researchers ever accept is 0.1 (10 %), requiring 96 plastic particles in samples, and would likely still allow significant differences between particle properties with the highest proportions. Increasing error reduction to be more robust, e.g., 0.01 (1 %), would require nearly 10,000 particles to be sampled. This suggests that high throughput techniques are required to advance certainty in microplastic concentrations (Primpke et al., 2020a, 2020b). These results bound estimates made by Kooi et al. (2021), who concluded that 500 plastic particles needed to be measured to estimate continuous distributions accurately, and estimates for minimum particle counts in other fields such as sediment (Eaton et al., 2019) and pollen (Chevalier et al., 2020).

3.4. Limitations

The major caveat to these results is that we must representatively and

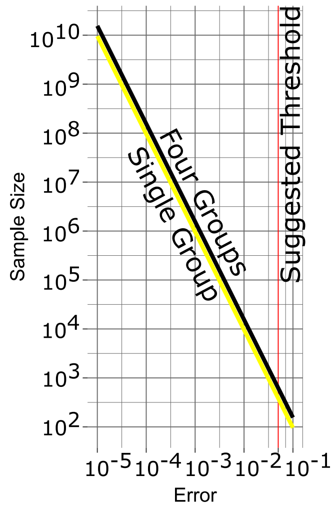

Fig. 4. The sample size needed to attain a error derived from the mathematical equations proposed in this study. The yellow line is the single group data while the black line is for samples with 4 groups. The red vertical line is at 0.05 and denotes our suggested minimum error. This figure can be used to quickly reference how many particles a sample should have when it is taken from a large population of plastic particles (> 100,000) e.g., a river, if it is to be representative. At the threshold proposed in this study of 0.05 error, there should be 384 particles in each sample when studying single groups, and for studying 4 groups, there should be 620 particles in the sample. Eqs. 1–3 were used to make this figure. The critical value was set to 1.96 and the expected probability was set to 0.5. All other parameters were varied as shown in the figure.

randomly subsample particles from environmental matrices and samples. Acute attention to the randomness is critical here. The particles must be randomly sampled from the entire distribution of particles for this equation to work correctly. Random sampling of particles will preserve the relative distributions of all particle properties (color, shape, size, polymer) and will be the best way for most study objectives.

As a field, we continue to learn how challenging representative and random sampling is given the diversity of plastic transport characteristics and plastic properties (Astray et al., 2023; Brander et al., 2020; Cid-Samamed and Diniz, 2023; Waldschläger et al., 2022). The tradeoff between reducing error and researcher effort or research cost is essential, and some samples may be impossible to reach the error threshold of 5 % we suggested (El Khatib et al., 2023). Range-finding exercises are beneficial to combine with the power analysis described in this work to assess the feasability of study questions and particle count goals.

The model proposed here does not include errors from subsampling, lab processing, or spectral analysis, which all impact the result of an analysis. This model only includes microplastic composition percentages and cannot be used to evaluate errors in estimating particle count-concentration (Parmar et al., 2023) or continuous distributions (Kooi et al., 2021). When presence-absence assessment of potentially rare polymeric materials is of interest, it is important to think beyond the error described here, which mainly tracks the highly abundant polymer types in a sample, and consider the likelihood that a particle will be observed in a sample given its abundance (Karlsson et al., 2020; Parmar et al., 2023). Future studies should address these limitations using appropriate a priori techniques and simulated rigorously to estimate complete error budgets (Morgado et al., 2022). Also, it is important to note that adding particle groups in Eq. (3) may lead to overly conservative sampling because properties may not be completely independent of one another, in cases where high correlation between groups is to be expected, Eq. (4) can be added. A complementary strategy to estimating a priori uncertainty is post hoc uncertainty. One such strategy is to collect subsamples until the uncertainty is at the desired limit based on the empirical uncertainty (Mintenig et al., 2020).

3.5. Next steps

We must apply this theory to new experimental conditions as the field develops new techniques (e.g., hyperspectral mapping instead of particle picking). For example, applying to particle picking is straightforward: pick all suspected microplastics and then randomly subsample the number of particles from Eqs. 1–3 based on the uncertainty. In spectral mapping, continue scanning greater amounts of the filter (selected randomly or based on substrate particle distribution) (Thaysen et al., 2020) until the number of microplastics necessary from Eqs. (1)–(3) has been reached. An in depth exploration into typical expectations for the correlations between groups should be explored, this would aid a priori analysis by reducing the sample count if correlations between groups can be confidently expected. Although large litter and macroplastic objects were not this manuscript's primary focus, we expect these equations to readily apply to the study of anthropogenic litter and hope that new advancements continue to be made to improve sample size estimation in that field (De Lange et al., 2023). We also want to point out that some freely available online calculators are beneficial for calculating sample sizes, confidence intervals, and other a priori metrics for estimating mean proportions (Brooks).

4. Conclusions

Our findings suggest that one can reliably (sub)sample and that the error between the subsample and the sample particle types will be determined, at least partially, by the initial and subsample counts. Researchers should use a priori statistical approaches to inform study design and assess study results based on particle counts in samples and subsamples. Subsampling savings in cost and time should generally be proportional to the proportion of subsample size of the sample size. We described a strategy to assess uncertainty in characterizing particle properties within a matrix and determine the minimum microplastic particle counts they should have in their samples based on their study goals. We provide helpful figures for researchers to use as a quick reference and suggest that no fewer than 96 microplastic particles be in samples or subsamples for quality data analysis and study design. This foundation provides grievance to the fact that many studies today have fewer particles in samples and, therefore individual samples likely poorly represent true microplastic properties. Last, we demonstrate that no more than 620 microplastic particles are needed for typical study designs in microplastic research today. This *a priori* approach is generally in agreement with findings from previous works. It establishes a statistical foundation for determining necessary microplastic counts in (sub)samples based on any desired uncertainty level.

CRediT authorship contribution statement

Laura Markley: Writing – review & editing, Writing – original draft, Visualization, Conceptualization. **Win Cowger:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Conceptualization. **Albert Koelmans:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Conceptualization. **Kshitij Upadhyay:** Writing – review & editing, Writing – original draft, Conceptualization. **Andrew B Gray:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Conceptualization. **Shelly Moore:** Writing – review & editing, Writing – original draft, Validation, Supervision, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

https://doi.org/10.5281/zenodo.10849581

Acknowledgments

Authors appreciate Hannah De Frond's early input on this work's direction. We acknowledge funding from McPike-Zima Foundation and the National Renewable Energy Laboratory which supported Win Cowger. Support for Laura Markley was provided by a NSF NRT award (Grant Award No. 1449617), a University fellowship and teaching assistantship through Syracuse University, and the Wen-Hsiung and Kuan-Ming Li Graduate Fellowship. Andrew Gray was supported in part by the USDA National Institute of Food and Agriculture Hatch Program [project number CA-R-ENS-5120-H] and the USDA Multistate Project 4170 [project number CA-R-ENS-5189-RR].

References

Astray, G., Soria-Lopez, A., Barreiro, E., Mejuto, J.C., Cid-Samamed, A., 2023. Machine learning to predict the adsorption capacity of microplastics. Nanomaterials 13, 1061. https://doi.org/10.3390/nano13061061.

Brander, S.M., Renick, V.C., Foley, M.M., Steele, C., Woo, M., Lusher, A., Carr, S., Helm, P., Box, C., Cherniak, S., Andrews, R.C., Rochman, C.M., 2020. Sampling and quality assurance and quality control: a guide for scientists investigating the occurrence of microplastics across matrices. Appl. Spectrosc. 74, 1099–1125.

Brandt, J., Fischer, F., Kanaki, E., Enders, K., Labrenz, M., Fischer, D., 2021. Assessment of subsampling strategies in microspectroscopy of environmental microplastic samples. Front. Environ. Sci. 8, 579676.

Brooks, S., Calculators. Select Statistical Consultants. URL ⟨https://select-statistics.co.uk/calculators/⟩ (accessed 9.15.23).

Chevalier, M., Davis, B.A.S., Heiri, O., Seppä, H., Chase, B.M., Gajewski, K., Lacourse, T., Telford, R.J., Finsinger, W., Guiot, J., Kühl, N., Maezumi, S.Y., Tipton, J.R., Carter, V.A., Brussel, T., Phelps, L.N., Dawson, A., Zanon, M., Vallé, F., Nolan, C., Mauri, A., de Vernal, A., Izumi, K., Holmström, L., Marsicek, J., Goring, S., Sommer, P.S., Chaput, M., Kupriyanov, D., 2020. Pollen-based climate reconstruction techniques for late Quaternary studies. Earth-Sci. Rev. 210, 103384 https://doi.org/10.1016/j.earscirev.2020.103384.

Cid-Samamed, A., Diniz, M.S., 2023. Recent advances in the aggregation behavior of nanoplastics in aquatic systems. IJMS 24, 13995. https://doi.org/10.3390/ijms241813995.

Cowger, W., Gray, A., Christiansen, S.H., DeFrond, H., Deshpande, A.D., Hemabessiere, L., Lee, E., Mill, L., Munno, K., Ossmann, B.E., Pittroff, M., Rochman, C., Sarau, G., Tarby, S., Primpke, S., 2020. Critical review of processing and classification techniques for images and spectra in microplastic research. Appl. Spectrosc. 74 https://doi.org/10.1177/0003702820929064.

Cowger, W., Gray, A.B., Guilinger, J.J., Fong, B., Waldschläger, K., 2021. Concentration depth profiles of microplastic particles in river flow and implications for surface sampling. Environ. Sci. Technol. 55, 6032–6041. https://doi.org/10.1021/acs.est.1c01768.

Cowger, W., Gray, A., Brownlee, S., Hapich, H., Deshpande, A., Waldschläger, K., 2022. Estimating floating macroplastic flux in the Santa Ana River, California. J. Hydrol.: Reg. Stud. 44, 101264 https://doi.org/10.1016/j.ejrh.2022.101264.

Cowger, W., Roscher, L., Jebens, H., Chamas, A., Maurer, B.D., Gehrke, L., Gerdts, G., Primpke, S., 2024. Generation of macro- and microplastic databases by high-throughput FTIR analysis with microplate readers. Anal. Bioanal. Chem. https://doi.org/10.1007/s00216-024-05127-w.

Daniel, W.W., Cross, C.L., 2018. Biostatistics: A Foundation for Analysis in the Health Sciences. Wiley.

De Frond, H., O'Brien, A.M., Rochman, C.M., 2023. Representative subsampling methods for the chemical identification of microplastic particles in environmental samples. Chemosphere 310, 136772. https://doi.org/10.1016/j.chemosphere.2022.136772.

De Lange, S.I., Mellink, Y., Vriend, P., Tasseron, P.F., Begemann, F., Hauk, R., Aalderink, H., Hamers, E., Jansson, P., Joosse, N., Löhr, A.J., Lotcheris, R., Schreyers, L., Vos, V., Van Emmerik, T.H.M., 2023. Sample size requirements for riverbank macrolitter characterization. Front. Water 4, 1085285. https://doi.org/10.3389/frwa.2022.1085285.

Dowle, M., Srinivasan, A., 2020. data.table: Extension of `data.frame`. URL ⟨https://rdatatable.gitlab.io/data.table/⟩. (Accessed 3.8.24).

Eaton, B.C., Moore, R.D., MacKenzie, L.G., 2019. Percentile-based grain size distribution analysis tools (GSDtools) – estimating confidence limits and hypothesis tests for comparing two samples. Earth Surf. Dyn. 7, 789–806. https://doi.org/10.5194/esurf-7-789-2019.

El Khatib, D., Langknecht, T.D., Cashman, M.A., Reiss, M., Somers, K., Allen, H., Ho, K.T., Burgess, R.M., 2023. Assessment of filter subsampling and extrapolation for quantifying microplastics in environmental samples using Raman spectroscopy. Mar. Pollut. Bull. 192, 115073 https://doi.org/10.1016/j.marpolbul.2023.115073.

Eriksen, M., Cowger, W., Erdle, L.M., Coffin, S., Villarrubia-Gómez, P., Moore, C.J., Carpenter, E.J., Day, R.H., Thiel, M., Wilcox, C., 2023. A growing plastic smog, now estimated to be over 170 trillion plastic particles afloat in the world's oceans—urgent solutions required. PLoS One 18, e0281596. https://doi.org/10.1371/journal.pone.0281596.

Erni-Cassola, G., Zadjelovic, V., Gibson, M.I., Christie-Oleza, J.A., 2019. Distribution of plastic polymer types in the marine environment; a meta-analysis. J. Hazard. Mater. 369, 691–698. https://doi.org/10.1016/j.jhazmat.2019.02.067.

European Commission. Joint Research Centre. Institute for Environment and Sustainability., MSFD Technical Subgroup on Marine Litter., 2013. Guidance on monitoring of marine litter in European seas. Publications Office, LU.

Guo, Z., Boeing, W.J., Xu, Y., Borgomeo, E., Mason, S.A., Zhu, Y.-G., 2021. Global meta-analysis of microplastic contamination in reservoirs with a novel framework. Water Res. 207, 117828 https://doi.org/10.1016/j.watres.2021.117828.

Karlsson, T.M., Kärrman, A., Rotander, A., Hassellöv, M., 2020. Comparison between manta trawl and in situ pump filtration methods, and guidance for visual identification of microplastics in surface waters. Environ. Sci. Pollut. Res 27, 5559–5571. https://doi.org/10.1007/s11356-019-07274-5.

Koelmans, A.A., Mohamed Nor, N.H., Hermsen, E., Kooi, M., Mintenig, S.M., De France, J., 2019. Microplastics in freshwaters and drinking water: Critical review and assessment of data quality. Water Res 155, 410–422.

Kooi, M., Koelmans, A.A., 2019. Simplifying microplastic via continuous probability distributions for size, shape, and density. Environ. Sci. Technol. Lett. 6, 551–557.

Kooi, M., Primpke, S., Mintenig, S.M., Lorenz, C., Gerdts, G., Koelmans, A.A., 2021. Characterizing the multidimensionality of microplastics across environmental compartments. Water Res, 117429.

Martí, E., Martin, C., Galli, M., Echevarría, F., Duarte, C.M., Cózar, A., 2020. The colors of the ocean plastics. Environ. Sci. Technol. 54, 6594–6601. https://doi.org/10.1021/acs.est.9b06400.

Miller, R.Z., Watts, A.J.R., Winslow, B.O., Galloway, T.S., Barrows, A.P.W., 2017. Mountains to the sea: River study of plastic and non-plastic microfiber pollution in the northeast USA. Mar. Pollut. Bull. 124, 245–251. https://doi.org/10.1016/j.marpolbul.2017.07.028.

Mintenig, S.M., Kooi, M., Erich, M.W., Primpke, S., Redondo- Hasselerharm, P.E., Dekker, S.C., Koelmans, A.A., Van Wezel, A.P., 2020. A systems approach to understand microplastic occurrence and variability in Dutch riverine surface waters. Water Res. 176, 115723 https://doi.org/10.1016/j.watres.2020.115723.

Moore, C.J., Lattin, G.L., Zellers, A.F., 2011. Quantity and type of plastic debris flowing from two urban rivers to coastal waters and beaches of Southern California. Rev. De. Gest. ão Coste Integr. 11, 65–73.

Morgado, V., Palma, C., Bettencourt da Silva, R.J.N., 2022. Bottom-up evaluation of the uncertainty of the quantification of microplastics contamination in sediment samples. Environ. Sci. Technol. 56, 11080–11090. https://doi.org/10.1021/acs.est.2c01828.

Nava, V., Chandra, S., Aherne, J., Alfonso, M.B., Antão-Geraldes, A.M., Attermeyer, K., Bao, R., Bartrons, M., Berger, S.A., Biernaczyk, M., Bissen, R., Brookes, J.D., Brown, D., Cañedo-Argüelles, M., Canle, M., Capelli, C., Carballeira, R., Cereijo, J.L., Chawchai, S., Christensen, S.T., Christoffersen, K.S., De Eyto, E., Delgado, J., Dornan, T.N., Doubek, J.P., Dusaucy, J., Erina, O., Ersoy, Z., Feuchtmayr, H., Frezzotti, M.L., Galafassi, S., Gateuille, D., Gonçalves, V., Grossart, H.-P., Hamilton, D.P., Harris, T.D., Kangur, K., Kankılıç, G.B., Kessler, R., Kiel, C., Krynak, E.M., Leiva-Presa, À., Lepori, F., Matias, M.G., Matsuzaki, S.S., McElarney, Y., Messyasz, B., Mitchell, M., Mlambo, M.C., Motitsoe, S.N., Nandini, S., Orlandi, V., Owens, C., Özkundakci, D., Pinnow, S., Pociecha, A., Raposeiro, P.M., Rõõm, E.-I., Rotta, F., Salmaso, N., Sarma, S.S.S., Sartirana, D., Scordo, F., Sibomana, C., Siewert, D., Stepanowska, K., Tavşanoğlu, Ü.N., Tereshina, M., Thompson, J., Tolotti, M., Valois, A., Verburg, P., Welsh, B., Wesolek, B., Weyhenmeyer, G.A., Wu, N., Zawisza, E., Zink, L., Leoni, B., 2023. Plastic debris in lakes and reservoirs. Nature 619, 317–322. https://doi.org/10.1038/s41586-023-06168-4.

Nel, H.A., Chetwynd, A.J., Kelleher, L., Lynch, I., Mansfield, I., Margenat, H., Onoja, S., Goldberg Oppenheimer, P., Sambrook Smith, G.H., Krause, S., 2021. Detection limits are central to improve reporting standards when using Nile red for microplastic quantification. Chemosphere 263, 127953. https://doi.org/10.1016/j.chemosphere.2020.127953.

Pabortsava, K., Lampitt, R.S., 2020. High concentrations of plastic hidden beneath the surface of the Atlantic Ocean. Nat. Commun. 11, 4073. https://doi.org/10.1038/s41467-020-17932-9.

Parmar, S., Arbuckle-Keil, G., Kumi, G., Fahrenfeld, N.L., 2023. Urban stormwater microplastic size distribution and impact of subsampling on polymer diversity. Environ. Sci.: Process. Impacts 25, 1374–1384. https://doi.org/10.1039/D3EM00172E.

Primpke, S., Godejohann, M., Gerdts, G., 2020b. Rapid identification and quantification of microplastics in the environment by quantum cascade laser-based hyperspectral infrared chemical imaging. Environ. Sci. Technol. 54, 15893–15903. https://doi.org/10.1021/acs.est.0c05722.

Primpke, S., Christiansen, S.H., Cowger, W., De Frond, H., Deshpande, A., Fischer, M., Holland, E.B., Meyns, M., O'Donnell, B.A., Ossmann, B.E., Pittroff, M., Sarau, G., Scholz-Böttcher, B.M., Wiggin, K.J., 2020a. Critical assessment of analytical methods for the harmonized and cost-efficient analysis of microplastics. Appl. Spectrosc. 74 https://doi.org/10.1177/0003702820921465.

Qian, N., Gao, X., Lang, X., Deng, H., Bratu, T.M., Chen, Q., Stapleton, P., Yan, B., Min, W., 2024. Rapid single-particle chemical imaging of nanoplastics by SRS microscopy. Proc. Natl. Acad. Sci. USA 121, e2300582121. https://doi.org/10.1073/pnas.2300582121.

Rochman, C.M., Brookson, C., Bikker, J., Djuric, N., Earn, A., Bucci, K., Athey, S., Huntington, A., McIlwraith, H., Munno, K., De Frond, H., Kolomijeca, A., Erdle, L., Grbic, J., Bayoumi, M., Borrelle, S.B., Wu, T., Santoro, S., Werbowski, L.M., Zhu, X., Giles, R.K., Hamilton, B.M., Thaysen, C., Kaura, A., Klasios, N., Ead, L., Kim, J., Sherlock, C., Ho, A., Hung, C., 2019. Rethinking microplastics as a diverse contaminant suite. Environ. Toxicol. Chem. 38, 703–711. https://doi.org/10.1002/etc.4371.

Roscher, L., Fehres, A., Reisel, L., Halbach, M., Scholz-Böttcher, B., Gerriets, M., Badewien, T.H., Shiravani, G., Wurpts, A., Primpke, S., Gerdts, G., 2021. Microplastic pollution in the Weser estuary and the German North Sea. Environ. Pollut. 288, 117681 https://doi.org/10.1016/j.envpol.2021.117681.

Roscher, L., Halbach, M., Nguyen, M.T., Hebeler, M., Luschtinetz, F., Scholz-Böttcher, B. M., Primpke, S., Gerdts, G., 2022. Microplastics in two German wastewater treatment plants: Year-long effluent analysis with FTIR and Py-GC/MS. Sci. Total Environ. 817, 152619 https://doi.org/10.1016/j.scitotenv.2021.152619.

RStudio Team, 2020. RStudio: Integrated Development Environment for R.

Sutton, R., 2019. Understanding Microplastic Levels, Pathways, and Transport in the San Francisco Bay Region. San Francisco Estuary Institute.

Thaysen, C., Munno, K., Hermabessiere, L., Rochman, C.M., 2020. Towards Raman automation for microplastics: developing strategies for particle adhesion and filter subsampling. Appl. Spectrosc. 74, 976–988. https://doi.org/10.1177/0003702820922900.

Vickerstaff, V., Omar, R.Z., Ambler, G., 2019. Methods to adjust for multiple comparisons in the analysis and sample size calculation of randomised controlled trials with multiple primary outcomes. BMC Med. Res. Method. 19, 129. https://doi.org/10.1186/s12874-019-0754-4.

Waldschläger, K., Brückner, M.Z.M., Carney Almroth, B., Hackney, C.R., Adyel, T.M., Alimi, O.S., Belontz, S.L., Cowger, W., Doyle, D., Gray, A., Kane, I., Kooi, M., Kramer, M., Lechthaler, S., Michie, L., Nordam, T., Pohl, F., Russell, C., Thit, A., Umar, W., Valero, D., Varrani, A., Warrier, A.K., Woodall, L.C., Wu, N., 2022. Learning from natural sediments to tackle microplastics challenges: a multidisciplinary perspective. Earth-Sci. Rev. 228, 104021 https://doi.org/10.1016/j.earscirev.2022.104021.

Weber, F., Kerpen, J., 2023. Underestimating microplastics? Quantification of the recovery rate of microplastic particles including sampling, sample preparation, subsampling, and detection using µ-Ramanspectroscopy. Anal. Bioanal. Chem. 415, 2963–2973. https://doi.org/10.1007/s00216-022-04447-z.

Werbowski, L.M., Gilbreath, A.N., Munno, K., Zhu, X., Grbic, J., Wu, T., Sutton, R., Sedlak, M.D., Deshpande, A.D., Rochman, C.M., 2021. Urban Stormwater Runoff: A Major Pathway for Anthropogenic Particles, Black Rubbery Fragments, and Other Types of Microplastics to Urban Receiving Waters. ACS EST Water 1, 1420–1428. https://doi.org/10.1021/acsestwater.1c00017.

Wickham, H., 2016. ggplot2: elegant graphics for data analysis. URL ⟨https://ggplot2-book.org/⟩. (accessed 3.8.24).

Wickham, H., Girlich, M., 2022. tidyr: Tidy Messy Data. URL ⟨https://tidyr.tidyverse.org/⟩. (accessed 3.8.24).

Wickham, H., François, R., Henry, L., Müller, K., 2020. dplyr: A Grammar of Data Manipulation. ⟨https://dplyr.tidyverse.org/⟩. (accessed 3.8.24).