Metadata for dataset "*Global data set on micro- and mesoplastic loads in marine sediments*"

Table 1. Description of the dataset

Title of dataset	*Global data set on micro- and mesoplastic loads in marine sediments*
URL of dataset	
Abstract	*In this dataset, we synthesize available estimates of micro- and mesoplastic stocks in marine sediments, from the upper intertidal to trenches, with the aim of estimating the global stock of plastic (10 µm – 2.5 cm) in ocean sediments. Indeed, while the global load of plastic in surface waters has been estimated and known to represent only a small fraction of the expected stock of plastic in the global ocean, no effort has been made to estimate the stock in marine sediments, despite they are believed to be the major sink of oceanic plastic. To fill this gap, we conducted a literature search and extracted estimates of plastic stocks in sediments available until March 2020, integrated with a search done in February 2021 to increase data for under-represented habitats. In this dataset, we report 1,649 estimates of plastics in marine sediments from 119 papers, including related information on sampling, sample processing and particle size and weight.*
Keywords	*Plastic; Sediment; Seafloor; Global*
Dataset lead author	*Cecilia Martin*
Position of data author	*Position during collection of data: graduate student*
Address of data author	*Address of the author during the collection of the data: 4700, KAUST, 23955 Thuwal, Saudi Arabia* *Current address: AlRaidah Digital City, Al Nakhil District 3807, Riyadh 12382 – 6726, Saudi Arabia*
Email address of data author	*Cecilia.martin@kaust.edu.sa*
Primary contact person for dataset	*Cecilia Martin*
Position of primary contact person	*Position during collection of data: graduate student*
Address of primary contact person	*Address of the author during the collection of the data: 4700, KAUST, 23955 Thuwal, Saudi Arabia* *Current address: AlRaidah Digital City, Al Nakhil District 3807, Riyadh 12382 – 6726, Saudi Arabia*
Email address of primary contact person	*Cecilia.martin@kaust.edu.sa*
Organization associated with the data	*King Abdullah University of Science and Technology (KAUST)*
Usage Rights	*Publicly available and free to use*
Geographic region	*Global study*

Geographic coverage	*Global coverage. Latitudinal range: -68º N to +85º N; longitudinal range: -177º E to +177º E*
Temporal coverage - Begin sampling date	*2004*
Temporal coverage - End sampling date	*2019*
General study design	*We searched for original studies reporting on plastic abundance in marine sediments in two of the main scientific search engines, Web of Science and Google Scholar. We screened the high throughput of literature obtained from the search in order to include only trustable (i.e., peer-reviewed) and relevant studies (i.e., effectively tackling plastics in marine sediments). From the studies kept after screening, we extracted data on plastic abundance, on sampling, on sample processing methodologies, on sample characteristics and on particle features. We compiled a dataset of 1,649 records of plastic abundance in sediments extracted from 119 original studies. To homogenize data, we converted all the abundance values to the same unit (number of items per kg of sediment and number of items per cm^3) and we corrected the values based on the lower size detection limit, different across studies, and on the recovery rate of the method use to extract plastics from sediments. Lastly, we separated the abundance of fibers from the abundance of all the other item types, named non-fibrous items, due to the uncertainty around the confirmation of the plastic nature of fibers and the high risk of contamination of the sample from airborne fibers during sample processing.*
Methods description	*We searched the literature for estimates of plastic (non-fibrous and plastic fiber items) abundance in marine sediments, ranging from intertidal sediments to deep sea trenches.* *Two literature searches were conducted on 5th March 2019 and on 10th March 2020 querying Web of ScienceTM using the terms "TS=(plastic AND marine AND sediments) NOT TS=(freshwater AND river)" and "TS=(Microplastic AND seafloor)". Two complementary searches were conducted in Google Scholar on the 7th May 2019 using the search terms "allintitle: microplastic sediments" and "allintitle: plastic sediments". We then inspected all the studies cited in the papers retrieved from the searches described above. These searches yielded a total of 962 published studies dating back to 1972. Studies were included based on meeting the following criteria: (a) the study was conducted in a marine habitat, extending from the deepest trenches to the upper intertidal zone; (b) plastic quantified was identified in sediment and (c) was not macroplastic (> 25 mm); and (d) data presented was primary (i.e. not reported from a different source) and peer-reviewed. On 26th February 2021, we did an additional search in Google Scholar focused on habitats for which the previous search generated only a few records (i.e. seagrass, ridges, rises and trenches), by using as keywords "microplastic" and "seagrass" or "deep sea" or "trench". With the additional research, we aimed at collating at least 50 records per habitat, however, the number of records obtained for trenches (26 records) is limited by the published research to date.* *The research yielded a total of 119 studies, from which the following data were extracted: site name and coordinates (decimal degrees); habitat type; water depth (m); sampling date; sampling method; solution used for the density separation; occurrence of chemical characterization (e.g. use of FTIR or Raman); occurrence of mitigation measures against contamination; the size of the smaller plastic*

	particle identified or the lower size fraction targeted (i.e., the lower detection limit); depth of the sample in the sediment and thickness (cm); the quantity of plastic particles identified (incl. ±SE/SD) per sampling site and number of replicates; relative frequency of fibers and non-fibrous plastics; mean weight (mg) and size of plastic particles (mm). When coordinates were not provided, the location name was input into Google Maps to generate site coordinates. However, when identifiable names were not provided, coordinates were generated by cross-referencing between maps provided in the publication to generate the best estimate of site location using Google maps. All site coordinates were then used to calculate individual distance between the sampling site and the nearest land mass using the measure tool on Google maps. If the sampling date was not provided in the paper or by the authors upon request (6 studies out of 119), we assumed that samples have been collected 2 years before publication, since the median lag between publication date and available sampling dates was 2 years. Sampling methods were pooled into four main categories (i.e. manual collection; corers; dredges and grabs; lander system). If the occurrence of a chemical characterization of the item was not specified, we assumed that it was not conducted. If the chemical characterization was conducted on all items except fibers, we specified it. When data provided was reported in a figure, the data extraction software GraphClickTM was used to extract it from plots where able.
Analytical methods	*All plastic stock quantities were converted to n kg^{-1} (sediment dry weight, DW) and to n cm^{-3} to homogenize data. To do so we multiplied or divided the abundance value per the dry bulk density. Since the dry bulk density was not given in the 119 studies, we assigned to each record of the dataset a dry bulk density value obtained from the literature. To do so, we found the closest site to each sampling point in our dataset, within the same depth range, for which a dry bulk density value is available from the literature. Alternatively, we identified sites for which Atwood et al., 2020 calculated the dry bulk density (DBD) by using a pedotransfer equation that derives DBD values from carbon content in sediment samples.* *For studies that reported total number of plastic items explicitly including synthetic fibers but pooled non-fibrous plastic and fibers in reporting estimates of plastic loads, we assumed that the fraction of fibers in the total pool of plastic items corresponded to the mean value (± SE) across studies of 0.61 ± 0.01 (N = 1097 estimates).* *In order to account for the variability due to the plastic extraction method, explicitly the use of different solutions across studies during the density separation step, we corrected the abundance values based on the recovery rates of the method applied. If the authors conducted a method validation, the recovery rates were extracted from the study itself. If the method validation was not conducted, we checked if a referring method was cited and we used the recovery rates published in the cited paper. Otherwise, we used recovery rates as those reported in the literature. Specifically, Quinn et al. (2017) estimated that 72% of the plastic is retrieved using distilled water and 89% using NaCl. The average recovery rate across references for use of NaI is 91.3% (Claessens et al., 2013; Crichton et al., 2017; Quinn et al., 2017) and for $CaCl_2$ is 81% (Crichton et al., 2017, Li et al., 2018). Coppock et al., 2017 and Imhof 2012 estimated recovery rates of 95.5% for $ZnCl_2$; Zhang et al., 2020 of 93% for NaH_2PO_4. The Potassium Formate recovery is not specified in literature (Mai et al., 2018), hence, being a solution with a density of 1.5 g cm^{-3}, we used, as recovery rate, the mean of the recovery rates of all the solutions with densities above 1.4 g cm^{-3} (i.e. 93.3%). Lithium*

	metatungstate has a recovery of 81% (Masura et al., 2015) and canola oil has a recovery of 89.3% (Lechthaler et al., 2020). *In order to account for the differences in the minimum plastic size assessed across studies, we corrected the abundance values based on the size distribution of particles. To build the distribution, we first extracted the size distribution of plastic particles from those studies that could provide a lower detection limit < 100 µm (N = 37 studies). We did not include studies that assessed only larger plastic sizes, > 100 µm, because they do not consider a large fraction of particles, as later confirmed by the size distribution. The size distribution in the included papers is reported as proportion of plastic in size bins. Since size bins provided in the papers have different lengths and boundaries, we normalized the value of proportion of plastic in each bin for the length of the bin and we defined new bins. Finally, we averaged the normalized values in each size bin across studies and we smoothed the distribution obtained. According to the model, sizes of plastics distribute in the marine sediments following an inverse power-law distribution, defined by the following equation:* $Ln(Y) = 0.92\ (\pm 0.03) \times Ln(X) – 3.8$ *Where Y is the normalized proportion of plastic particles, X is the central point of the size bin and 0.92 (±0.03) and -3.8 are the estimates of the slope and the intercept, respectively.* *The inverse power-law size distribution of plastic particles was used to model the cumulative distribution of plastic particles in sizes (i.e., $Ln(Y) = 1.058\ (\pm 0.08) \times Ln(X) – 2.4$), from which we calculated correction factors to account for the underestimation of plastic abundance obtained when only large sizes are considered. For instance, when only particles >100 µm, >500 µm or >1000 µm are extracted from sediments, the abundance is 11, 63 and 131 times underestimated, respectively, compared to when particles down to 10 µm are considered. By multiplying the plastic abundance values for the correction factors (that we indicate in the dataset), we obtain an estimate of the abundance of plastics ≥10 µm in all samples. When the lower size detection limit was not reported in a paper, we did not apply any correction factor, to avoid inflating the estimate.*
Quality control	*Despite the dataset was compiled by 3 authors, the compilation was supervised by one author to ensure consistency across data collectors.* *Moreover, the dataset was double checked by the dataset lead author to identify possible mistakes during collection, for example:* - *Coordinates were represented in a map to ensure that all points were located in seawater or along the coast. Particular attention was given to points located in areas of the world where sampling is challenging (e.g., the open ocean, the Arctic and Antarctic), which coordinates were double checked.* - *If stocks of plastics were larger or smaller than the 95% CI around the mean estimate, they were double checked to ensure that they were correctly reported.* - *The categories of each variable were checked in R studio, by using the code levels(), to ensure the name consistency.* - *Consistency between the water depth and the habitat type was also checked.*

Table 2. Description of the variables (i.e., columns) in the "Dataset" tab

Column name	Definition	Units
ID	Number assigned to each record	NA (Integer number)
Reference	The number (from 1 to 119) corresponds to the number assigned to the original study in the tab "References" (Table 3).	NA (Integer number)
Continent	The continent where the sample was collected. If the sample was collected outside the territorial sea of a continent, we either indicated EEZ (Exclusive Economic Zone) or High Seas, depending on the distance to the land.	NA (Text)
Latitude	Latitude of sampling point.	ºN
Longitude	Longitude of sampling point.	ºE
Distance_land	The distance from the sampling point and the closest land mass, measured in Google Maps.	Miles
Habitat_Type	Biome of the sampling point. It includes (in alphabetical order): abyss; beach (i.e., intertidal sediments); continental shelf; estuary; harbour; lagoon; mangroves; mudflat; reef; ridge; seagrass; slope; trench.	NA (Text)
Water_Depth	The water depth of the sampling point as reported in the original study or extracted from GEBCO (2020) database (doi:10.5285/a29c5465-b138-234d-e053-6c86abc040b9). Those extracted from GEBCO are highlighted in orange.	meters
Mean Population in 110 km radius	Mean population density in a radius of 110 km from the sampling point, extracted from CIESIN, Gridded Population of the World, Version 4 (GPWv4), Population Density, Revision 11 (2018).	Number of people per km^2
Sampling_date	The date of the sample collection as reported in the original study or from the authors of the paper upon request. However, dates highlighted in yellow were not provided and were therefore estimated from the median lag between sampling and publication date of the study with known sampling date.	NA (Integer number)
Sampling_method	The methods of sample collection were categorized in 4 classes: corers; dredges & grabs; lander system; manual collection.	NA (Text)
Lower_limit (µm)	Lower detection size limit of plastic particles extracted from the sample(s). It is either the smallest particles found, or the smallest filter used to separate the sample.	micrometers
Correction_factor	This correction factor is calculated from the cumulative distribution of plastic particles in sizes (see above "Analytical methods"), as follows: Correction_factor=0.0915*(Lower_limit)$^{-1.058}$ This factor is used to correct the abundance value (see later "Corrected_Fibers (n kg-1)" and "Corrected_Non-fibrous (n kg-1)").	NA (decimal number)
Upper_limit (mm)	Upper detection size limit of plastic particles	millimeters

	extracted from the sample(s). It is either the largest particles found, or the largest filter used to separate the sample.	
Density_separation_method	Solutions used to extract plastics from sediments through density separation.	NA (Text)
Recovery_density_separation (%)	Recovery rates of the extraction method used (see above "Analytical methods").	%
Screening	Method used to screen the sample(s) after density separation: either visually with microscope or visually without microscope.	NA (Text)
FTIR_Raman	Occurrence of a chemical characterization of the plastic particles extracted. The two most common methods are Fourier Transformed Infrared Spectroscopy (FTIR) and Raman, but if other chemical characterization methods were used, we still indicated "YES" in this column.	NA (Text)
Mitigation	Occurrence of mitigation measures against contamination during sample processing (e.g., clean thoroughly material with distilled water, wear a cotton lab coat and/or note clothes worn during sample processing, work in a clean environment, use controls and blanks).	NA (Text)
Depth_min (cm)	Upper depth of the sediment sample(s). If, for example, a sample is collected in the top 5 cm of sediment, this value is 0.	centimeters
Depth_max (cm)	Lower depth of the sediment sample(s). If, for example, a sample is collected in the top 5 cm of sediment, this value is 5.	centimeters
Central_depth (cm)	Central depth of the sediment sample(s). If, for example, a sample is collected in the top 5 cm of sediment, this value is 2.5.	centimeters
Thickness (cm)	Difference between the Depth_max and the Depth_min value.	centimeters
Dry Bulk Density (DBD)	Value of dry bulk density (DBD) of the closest site to the sampling point, within the same depth range. Values of DBD were either obtained from the literature or calculated by Atwood *et al*., 2020, by using a pedotransfer equation that derives DBD values from carbon content in sediment samples. Atwood, T. B., Witt, A., Mayorga, J., Hammill, E., & Sala, E. (2020). Global patterns in marine sediment carbon stocks. *Frontiers in Marine Science*, *7*, 165.	g cm^{-3}
OriginalSource_DBD	Literature reference of the original source for the values of DBD reported in the column "Dry Bulk Density (DBD) or for the values of carbon content used in Atwood *et al.*, 2020 to calculate the dry bulk density values.	NA (Text)
DataSource_DBD	Literature reference of the datasets from which DBD values or carbon content values were extracted.	NA (Text)
Sample_Number	Number of samples processed for each record.	NA (Integer number)
Quantity	Plastic abundance in the sample(s).	The unit is specified in the column "Units"
(+/-)	Error of the estimate reported in "Quantity", if more	The unit is

	than one sample was processed for each record.	specified in the column "Units"
SD/SE	Type of error (e.g., SD, SE, 95% CI).	NA (Text)
Units	Unit of the plastic abundance estimate in "Quantity" (e.g., number of items per kg of sediment or n kg^{-1}, n l^{-1}, n m^{-2}, n m^{-3}).	NA (Text)
Abundance_converted (n kg-1)	Abundance of total plastic (both fiber and non-fibrous) in number of items per kg of sediment. To convert abundance values that were reported as number of items per unit volume, we divided the abundance per the value of DBD.	Number of items per kilogram of sediment
Abundance_fibers (n kg-1)	Abundance of fibers (in number of fibers per kg of sediment), if provided by the original study. If not directly provided by the study, this value was calculated by multiplying the abundance of total plastic (value in "Abundance_converted") for the value of proportion of fibers over total number of plastics provided by the original study (value in "Proportion").	Number of items per kilogram of sediment
Abundance_Non-fibrous (n kg-1)	Abundance of all type of items that are not fibers (e.g., hard fragments, films, foams, pellets, beads) in number of non-fibrous plastic per kg of sediment, if provided by the original study. If not directly provided by the study, this value was calculated by multiplying the abundance of total plastic (value in "Abundance_converted") for the value of proportion of non-fibrous plastics over total number of plastics provided by the original study (value in "Proportion").	Number of items per kilogram of sediment
Proportion	Proportion of fibers over the total number of plastics extracted from the sample(s) as reported in the original study.	NA (Decimal number)
Estimated_proportion	Estimated proportion of fibers over the total number of plastics extracted from the sample(s). If the proportion value was given in the original study, the value we report in this column is the same as in the "Proportion" column, otherwise we assigned the value of 0.6, which is the mean value across 1097 studies that provided the proportion value.	NA (Decimal number)
Estimated_Fibers (n kg-1)	Estimated abundance of fibers in the sample(s). If available, the value was left equal to the one in the "Abundance_fibers" column. Otherwise, we assigned a value calculated by multiplying the abundance of total plastic (value in "Abundance_converted") per proportion of fibers in the sample (value in "Estimated_proportion").	Number of items per kilogram of sediment
Estimated_Non-fibrous (n kg-1)	Estimated abundance of fibers in the sample(s). If available, the value was left equal to the one in the "Abundance_fibers" column. Otherwise, we assigned a value calculated by multiplying the abundance of total plastic (value in "Abundance_converted") per proportion of non-fibrous plastics in the sample (1 - value in "Estimated_proportion").	Number of items per kilogram of sediment
Corrected_Fibers (n kg-1)	This value was obtained multiplying the value in "Estimated_Fibers" per the value in "Correction_factor" and divided by the value in "Recovery_Density_Separation"), in order to correct	Number of items per kilogram of sediment

	the abundance of fibers per the underestimation given by the size fraction of the plastic tackled and the recovery rate of the solution used for the density separation.	
Corrected_Non-fibrous (n kg-1)	This value was obtained multiplying the value in “Estimated_Non-fibrous” per the value in “Correction_factor” and divided by the value in “Recovery_density_separation”), in order to correct the abundance of non-fibrous items per the underestimation given by the size fraction of the plastic tackled and the recovery rate of the solution used for the density separation.	Number of items per kilogram of sediment
Corrected_Fibers (n cm-3)	This value was obtained by converting the value in “Corrected_Fibers (n kg-1)”. To do the conversion, we multiplied the value of corrected fibers (n kg-1) per the DBD and divided by 1000	Number of items per cubic cm
Corrected_Non-Fibrous (n cm-3)	This value was obtained by multiplying the value of “Corrected_Non-fibrous (n kg-1)” per the DBD and dividing by 1000	Number of items per cubic cm
Sediment_accretion_rate	For seagrass and mangroves, we report the global estimate obtained from the literature (Duarte et al, 2013). For the other habitats, we derived the sediment accretion rate by the water depth using the equation in Middleburg et al., 1997. Duarte, C. M., Losada, I. J., Hendriks, I. E., Mazarrasa, I., & Marbà, N. (2013). The role of coastal plant communities for climate change mitigation and adaptation. *Nature Climate Change*, *3*(11), 961-968. Middelburg, J. J., Soetaert, K., & Herman, P. M. (1997). Empirical relationships for use in global diagenetic models. *Deep Sea Research Part I: Oceanographic Research Papers*, *44*(2), 327-344.	cm y^{-1}

Table 3. Description of the variables (i.e., columns) in the “References” tab. This tab includes the details of the 119 studies from which the data in the “Dataset” tab were extracted or derived.

Column name	**Definition**	**Units**
Number	We assigned a number to each of the studies included in the dataset. The number assigned is casual, solely based on the chronological order of the data extraction.	NA (Integer number)
Author	List of authors of the original study.	NA (Text)
Year	Year of publication of the original study.	NA (Integer number)
Title	Title of the original study.	NA (Text)
Journal	Journal where the original study is published.	NA (Text)
DOI	Digital Object Identifier of the original study.	NA (Text)
Link	Hyperlink to the original study.	NA (Text)

Table 4. Description of the variables (i.e., columns) in the "References_DBD" tab. This tab includes the details of the publications or datasets from which the values in the "Dry Bulk Density (DBD)" column of the "Dataset" tab were extracted or derived.

Column name	Definition	Units
Author	List of authors	NA (Text)
Year	Year of publication of the original study.	NA (Integer number)
Title	Title of the original study.	NA (Text)
Journal/Database	Journal where the original study is published.	NA (Text)
Link	Hyperlink to the original study.	NA (Text)

Table 5. Description of the variables (i.e., columns) in the "Individual mass of particles" tab. This tab includes data on the individual mass of plastic particles extracted from those paper that provided this kind of values.

Column name	Definition	Units
Reference	The number corresponds to the number assigned to the original study in the tab "References" (Table 3).	NA (Integer number)
Central size (mm)	Size of the plastic particle or central value of the size range of the plastic particle(s).	millimeters
Weight (mg)	Mass of the plastic particle(s).	milligrams