Coordinated Environmental Monitoring Programme (CEMP) guidelines for the monitoring and assessment of seafloor litter

OSPAR Agreement 2017-06[1]

Source: EIHA 2023 Annex08

OSPAR Theme: Human activities

OSPAR descriptor: Marine litter

Indicator full name: Spatial and temporal trends of litter on the seabed collected during demersal trawls

COMMISSION DECISION (EU) 2017/848 :

Descriptor: 10 - Marine Litter

Criterion: D10C1 - The composition, amount and spatial distribution of litter on the coastline, in the surface layer of the water column, and on the seabed, are at levels that do not cause harm to the coastal and marine environment

Specifications and standardised methods for monitoring and assessment: For D10C1:
The composition, amount and spatial distribution of litter (excluding micro-litter) on the coastline and may additionally be monitored in the surface layer of the water column and on the seabed. Information on the source and pathway of the litter shall be collected, where feasible.

[1] English only. Update 2023

Authors: Josie Russell & Jon Barry

Contributors: Seafloor litter Expert Group (SLEG)

Contents

List of Abbreviations

Relating to OSPAR

OSPAR: Oslo – Paris convention

NAES: North-East Atlantic Environment Strategy

EIHA: Environmental Impacts of Human Activities

ICGML: Intersessional Correspondence Group on Marine Litter

SLEG: Seafloor Litter Expert Group

CEMP: Coordinated Environmental Monitoring Programme

JAMP: Joint Assessment and Monitoring Programme

GNS: Greater North Sea

CS: Celtic Seas

BB: Bay of Biscay and Iberian Coast

Others

IBTS: International Bottom Trawl Surveys

DOI: Digital Object Identifier

TV: Threshold Value

GES: Good Environmental Status

ICES: International Council for the Exploration of the Sea

MSFD: Marine Strategy Framework Directive

TGML: Technical Group on Marine Litter

QA/ QC: Quality Assurance/ Quality Check

1. Introduction

OSPAR is the mechanism by which 15 Governments and the EU cooperate to protect the marine environment of the North-East Atlantic. The OSPAR maritime area is divided into five regions, the Arctic Waters, Greater North Sea (GNS), Celtic Seas, Bay of Biscay and Iberian Coast, and the Wider Atlantic (Figure 1). Since 1972, contracting parties to the OSPAR convention have worked together to identify threats to the marine environment and to put in place programmes and measures to ensure effective collective and national action to combat them. Under the North-East Atlantic Environment Strategy (NEAES) 2020-2030, OSPAR has strategic objectives to significantly reduce marine litter to levels that do not cause adverse impacts. These levels, referred to as "Good Environmental Status" (GES), are also the objectives set by the European Union (EU) in the Marine Strategy Framework Directive (MSFD, 2008/56/EC).

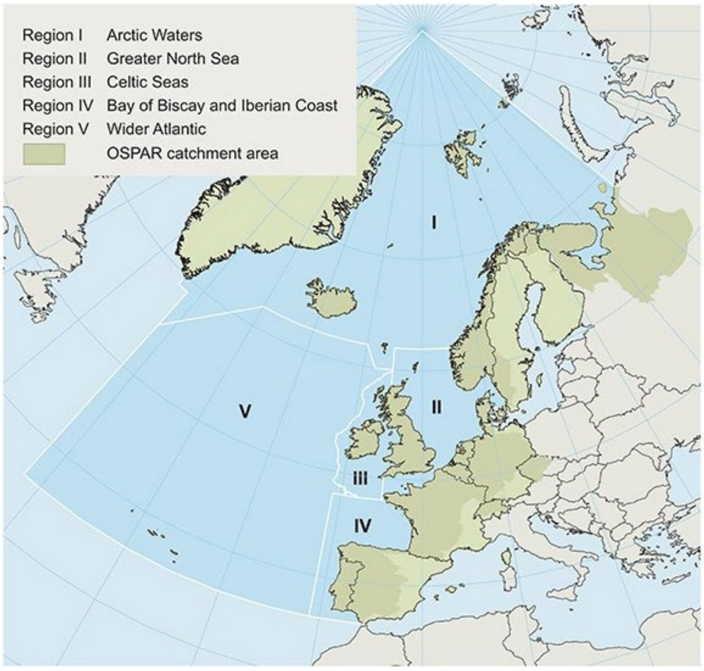

Figure 1: Map showing OSPAR regions.

Measures to tackle marine litter by reducing its input and removing it from the marine environment are implemented at the European level through the provisions resulting from the MSFD and other relevant directives such as the SUPD, the PRF-D and others and, at the regional OSPAR level through the OSPAR Regional Action Plan on Marine Litter (RAP ML) and at the national level through countries own policy and legislation.

OSPAR's work on Marine Litter is coordinated through the Intersessional Correspondence Group on Marine Litter (ICGML) which is a subsidiary group from the Environmental Impacts of Human Activities Committee (EIHA). Its Coordinated Environmental Monitoring Programme (CEMP) aims to deliver comparable data across the OSPAR area. It is hoped that assessments using this data can address specific questions raised in OSPAR's Joint Assessment and Monitoring Programme (JAMP).

Indicators have been developed to measure actions against marine litter and assess their effectiveness in reducing marine litter according to specific criteria to achieve GES. OSPAR currently assesses beach litter, litter on the seabed, plastic particles in fulmar stomachs and litter ingested by sea turtles, and is working on a new indicator for microplastics in sediments as part of its monitoring programme. The programme aims for the abundance, trends, and composition of marine litter in the OSPAR Maritime Area to be determined for different environmental compartments (floating, seabed and coast).

Litter on the seafloor (>2.5cm) has been collected and recorded since as early as 1992 by scientists monitoring fish stocks as part of ICES Bottom Trawl Surveys (Galgani et al., 2000). The litter enters the trawls, alongside the fish and is counted, weighed, and classified as an added value to the fisheries survey. The main challenge for the monitoring of litter on the seabed is that there are differences among fisheries surveys in terms of aims, timing, locations, and gear types. This complicates further marine litter analysis and needs to be taken into consideration for future assessments. A better understanding of litter lifecycle and factors which influence its rate of burying or dispersal is key to understanding what the data means.

It is important to recognise that the current monitoring programme relies on bottom trawling, for which the future is uncertain due to the environmental impact. Therefore, it is valuable to, whilst continuing with the current monitoring, also investigate other state of the art methods of monitoring seabed, which can be considered for future monitoring and to inform management strategies. Methods, such as using underwater camera systems, might be an important complement to increase our knowledge and understanding of litter occurrences in areas that are not trawled (i.e. rocky and/or deep) and to ensure that these habitats are not deposition areas for marine litter.

The aims of this document are to provide OSPAR Contracting Parties guidelines to set out details of the agreed monitoring and assessment approaches for litter on the seabed, including the approaches to be applied to realise coordination of monitoring and assessment. This document will introduce i) how to monitor litter on the seabed using trawling surveys, and ii) recommend methods to make an assessment on the data produced considering types of litter, spatial patterns, temporal changes, and trends in the distribution.

The ICES Working Group on Marine Litter (ICES WGML) has published technical guidelines, a photo guide and improved instructions for QA/QC, so rather than duplicate this work, the ICES documents are linked and referenced in this document (ICES, 2022).

The CEMP guidelines have been adapted to align with the European Commission's requirements for MSFD reporting and are available in the OSPAR Assessment Portal. The purpose is to assist those Contracting Parties that are EU member states in their national reporting commitments for MSFD Article 11 Monitoring Programmes.

2. Monitoring

2.1 Aims

The main aim of the OSPAR seafloor litter monitoring programme is to provide information about the spatial and temporal trends, and types and quantities of seafloor litter collected by the demersal trawl surveys. This is used to feed into the reporting of the quality status of the OSPAR area and to determine whether GES is achieved. This can also help to determine, and provide information on, how effective regional and national measures are at reducing marine litter. As litter is transported by currents, it is worth noting that levels are not necessary connected to measures applied in specific countries, therefore, it is important to consider results from a regional perspective.

The monitoring programme is producing seafloor litter data and metadata which are comparable across the OSPAR region. A standardised methodology to achieve harmonisation is based on the monitoring guidance of the MSFD Technical Expert Group on Marine Litter, which states that the best way to survey seafloor litter is by utilising existing trawl surveys. The ICES Manual for Seafloor Litter Data Collection and Reporting from Demersal Trawl Samples has also been produced to ensure harmonisation of the data collection (ICES, 2022).

Marine litter monitoring and assessment can also increase public awareness regarding the problem of marine litter. The trawls themselves also remove marine litter from the seafloor, and surveying the marine litter adds value to existing trawl surveys which already aim to quantify marine fisheries resources.

2.2 Quantitative objectives

As stated in the 2012 OSPAR Advice document on marine litter, target setting is difficult for almost all indicators under the MSFD Descriptor 10 for several reasons. The relationship between the types and amounts of marine litter in the environment and the degree of 'harm' caused at a population level, and in some cases individual level, is not fully understood. Threshold values are set to show the level at which there is no harm to the environment, but as there is not enough data to understand the relationship between harm and number of items of seafloor litter, this cannot be set. Sometimes threshold values can be set through expert judgement, so this may later be considered by the Seafloor Litter Expert Group (SLEG) if deemed appropriate.

2.3 Monitoring strategy

Most currently occurring trawl surveys suited for additional marine litter data collection are based on fisheries objectives and thus may not be optimal for measuring seafloor litter, while also running the risk that they may be discontinued or changed for reasons outside the scope of OSPAR litter monitoring. Initiating a targeted marine litter survey programme for higher resolution temporal trend monitoring would involve additional costs which would depend on the magnitude of the time trend to be detected with a level of statistical certainty and the time period within which it is to be found. Such additional trawling of the seafloor would also result in additional ecological damage. These additional costs and damage potentially outstrip the benefits of additional trawled surveys only focused on marine litter.

Contracting Parties currently have monitoring programmes in place at differing stages of implementation and, although there has been a lot of work done to improve and ensure harmonisation of the methodology and adequate QC of the data through the SLEG and ICES WGML, further work is still needed, especially to harmonise count data so it can be used in the assessments.

It is recommended that seafloor litter is monitored as an added value element of existing trawling surveys to monitor fish stocks. Programmes to monitor marine litter on the seafloor using existing trawling surveys have been introduced by several contracting parties, which led to the inclusion of seafloor litter in the OSPAR Common Indicator process. Almost all Contracting Parties organise fisheries surveys using bottom trawls across the OSPAR region, often in overlapping areas (especially in Region II, III, IV) (Figure 1).

2.4 Sampling strategy

Bottom trawl surveys are widely used for monitoring demersal stocks, therefore, site selection, sampling units (gear type and haul length) and survey periods are all determined by the fisheries programme. Fisheries surveys generally perform a number of test trawls and studies over the years to test and monitor catchability of different fish species and gears at different conditions. However, there are only a few such studies addressing the catchability of litter (O'Donoghue & Van Hal, 2018; Barry and Russell, 2022). Fisheries surveys have remained fairly constant over time in terms of sampling procedures and areas sampled. The litter data they generate therefore allows us to set baselines and understand trends based on sound methods in an environmental matrix where information normally comes at high cost.

Given that it is the fisheries surveys that largely are driving the sampling, the marine litter program and assessment method will need to be flexible. As a result of different monitoring strategies and sampling procedures, the amounts of litter collected will vary on different trawl surveys. For example, different gears will have different abilities to catch litter, which may additionally vary due to properties of the substrate. It would be highly desirable if parallel or alternate hauls could be carried out to estimate correction factors for comparison of marine litter data. This is not something which would easily happen as it would be costly and involve setting up trawls just for collection of litter. In the absence of such correction factors, mixing of survey data need to be approached with caution (e.g. GOV and BAK assessment). One approach, recommended below, is that the data are modelled, and the output standardised to a particular gear type.

Within the OSPAR region, most fish trawl surveys are coordinated by ICES expert groups. The survey data cover the Baltic Sea, Skagerrak, Kattegat, North Sea, English Channel, Celtic Sea, Irish Sea, Bay of Biscay and the eastern Atlantic from the Shetlands to Gibraltar.

2.5. Sampling methodology

For guidelines on seafloor litter sampling as part of trawl surveys, please follow the ICES Manual for Seafloor Litter Data Collection and Reporting from Demersal Trawl Samples (ICES, 2022). These have been designed to provide user friendly instructions for ship's crew and researchers collecting and recording seafloor litter.

2.6 Quality Assurance/ Quality Control

It is important that the data generated is of acceptable quality, Considerable effort has gone into ensuring that the data is consistent across trawl surveys and across reporting countries. To ensure the quality and integrity of seafloor litter monitoring data, the use of quality control and assurance measures, such as training for operators and use of pictorial guides of litter items needs to be introduced. A picture guide and sampling recommendations were developed by the ICES WGML (ICES, 2022). ICES WGML has also developed a web test for litter registration that was launched at the end of November 2022. Further development of QA/QC procedures for seabed litter form part of the future Terms of Reference for the ICES WGML.

2.7 Data reporting, handling, and management

DATRAS (the Database of Trawl Surveys) has been developed to collate and document survey data, assure data quality, standardise data formats and calculations, and ease data handling and availability. Those existing frameworks and databases have been adapted by the ICES Data Centre to accommodate OSPAR seafloor litter data.

Seafloor litter is submitted annually by Member States to the ICES DATRAS database (and can be downloaded via the data centre). This application has a download capability where csv files of the data points can be downloaded along with accompanying meta data. Additional columns have been added to filter data by OSPAR area, MSFD regions, EEZ and territorial waters. This application also offers standard web services which can be used by statistical packages, i.e. R scripting. Additionally, an overview of the submission status of all available seafloor litter data is also available. When downloading data users have the choice between using "Litter Exchange Data" or "Litter Assessment Output". We would recommend the latter because this outputs a single file linking both the haul data (e.g., location, length) and the litter data (e.g., type of litter and quantities found). As of 2020, there are 12 sets of data from the OSPAR region which can be downloaded. These correspond to 12 bottom trawl surveys. These are shown in table 1. Figure 2 shows the locations of the surveys.

Further information on data submission and extraction can be found in the ICES manual for seafloor litter data collection and reporting from demersal trawls (ICES 2022).

3. Assessment

3.1 Aims

The assessment method uses data from the OSPAR seafloor litter monitoring programme to assess the abundance, composition, and trends of litter on the seafloor collected during trawl surveys in the OSPAR region. One of the aims is to be able to produce assessments which show trends over space and over time. This information can be used to inform new actions on marine litter, to look at whether the litter reduction aims are being met and assess how effective certain measures which have been put in place at OSPAR or EU level are.

There are currently no agreed Threshold Values (TV) or Intermediate Measurable Targets (IMT) for seafloor litter at EU, OSPAR or national level, however this is something that should considered in the future. There are currently agreed TV values at EU level for beach litter and plastic in Fulmar stomachs, The fulmar TV, which is based on the paper of Van Franeker et al. (2021), has already been adopted by OSPAR as a regional threshold value and the beach litter TV is being considered by OSPAR.

In Europe and at OSPAR level, there have been numerous measures (policy and legislation) put in place to reduce marine litter. It is relevant to monitor and assess if specific measures lead to the desired reduction (decreasing trend) of specific litter. Relevant examples for seafloor litter are the EU Single Use Plastic and fishing gears directive (EU 2019/904) and the EU plastic bag directive (EU 2015/720). These policy groups should be harmonised across indicators.

3.2 Method

The assessment method below is based on the latest OSPAR intermediate assessment (Barry and Russell et al, 2022) which is part of the 2023 OSPAR Quality Status Report.

We recommend using data only from 2012 onwards as this was when the MSFD TGML set out the first set of guidelines for monitoring seabed litter as part of fisheries surveys, so we expect there to be reasonable consistency in terms of classification of litter. Before this date, the data is sparse and there was no harmonised protocol.

3.3 Area

The surveys mainly take place in the following three OSPAR regions Greater North Sea (GNS), Celtic Seas (CS) and Bay of Biscay and Iberian Coast (BB) (see also map provided as Figure 1). There are some sampling locations in the Wider Atlantic and also the Arctic Waters, however these areas have too little data, and the coverage is insufficient to make an assessment on. This can be checked for future assessments.

Table 1*: Survey programmes and gear types included in the seafloor litter assessment. Also shown are explanations of the fishing gear code abbreviations.*

Survey programme	Survey code	Type of gear used	Gear code
Beam Trawl Survey	BTS	Beam Trawl 4, 7 and 8 m	BT4A, BT4AI, BT7, BT8
French Southern Atlantic Bottom Trawl Survey	EVHOE	Grand Ouverture Verticale Trawl	GOV
French Channel Ground Fish Survey	FR-CGFS	Grand Ouverture Verticale Trawl	GOV
Irish Ground Fish Survey	IE-IGFS	Grand Ouverture Verticale Trawl	GOV
North Sea International Bottom Trawl Survey	NS-IBTS	Grand Ouverture Verticale Trawl	GOV
Portuguese International Bottom Trawl Survey	PT-IBTS	Norwegian Campell Trawl 1800/96	NCT
Scottish Rockall Survey	SCOROC	Grand Ouverture Verticale Trawl	GOV
Scottish West Coast Groundfish Survey	SCOWCGFS	Grand Ouverture Verticale Trawl	GOV
Spanish Gulf of Cadiz Bottom Trawl Survey	SP-ARSA	Baka Trawl	BAK
Spanish North Coast Bottom Trawl Survey	SP-NORTH	Baka Trawl	BAK
Demersal Young Fish Survey	DYFS	Beam Trawl 6 m	BT6
Spanish Porcupine Bottom Trawl Survey	SP-PORC	Porcupine Baka	PORB

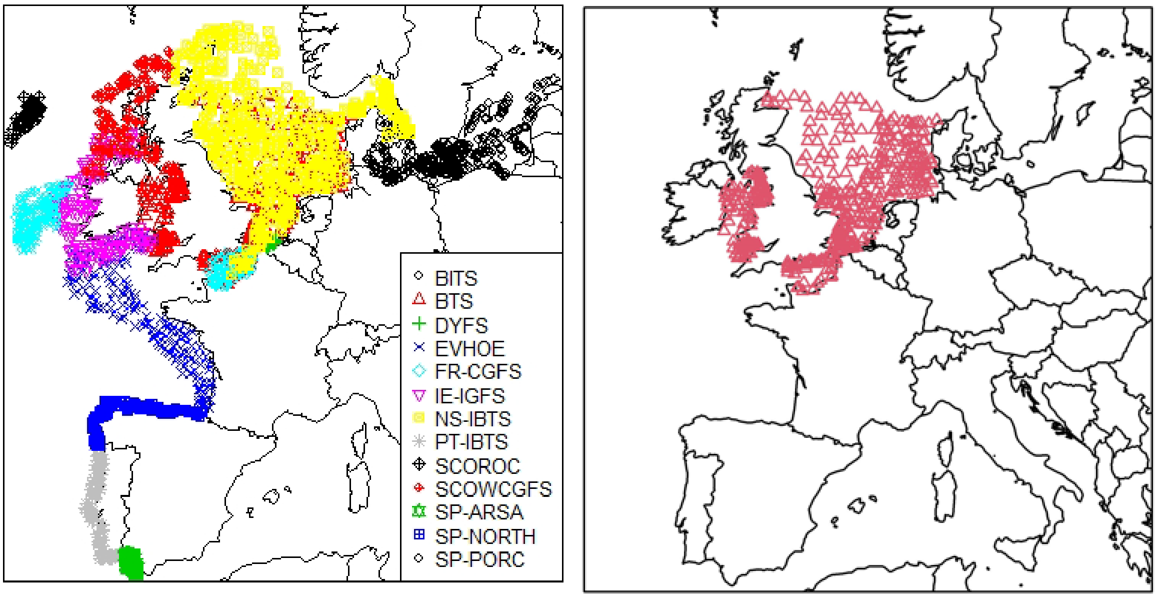

Figure 2: *Survey locations in 2018. Note that the PT-IBTS survey did not take place in 2018 and therefore its location is shown from 2016. BTS survey locations only are shown on the right due to overlap with other surveys. Note a survey outside OSPAR area is also included in the figure but not included in assessments.*

3.4 Considering variables

It is important that litter data values are not biased because of.

- The type of gear
- Unequal spatial sampling
- The area swept of the trawl

For gear type, informal studies have shown that, for the same areas, different types of gear can differ in catch by a factor of twelve in the counts of litter recorded (Barry and Russell et al., 2022). Weight of the trawl controls how deep into the sediment the gear will dig, and gear width (wing spread) controls the area that is swept. The number of litter items found in a trawl should be directly proportional to the area swept. Current knowledge suggests that GOV trawls capture only around 5% of the items on the seafloor and that actual numbers of seafloor litter are substantially higher than reported (O'Donoghue and Van Hal, 2018). Therefore, for example, the absence of litter in a haul does not mean that there is zero litter on the seafloor. This should be made clear in any assessment of the data. Future assessments should consider again how to take into account the differences in litter recoveries between gear types, as more evidence is provided to understand. We know methods such as using beam trawls have shown higher recoveries (Barry and Russell et al., 2022).

Unequal sampling effort in space is important because if one part of a region is sampled more heavily than another part and we take a simple mean of all the results, our mean will be biased towards the area with the greater sampling effort. This is remedied in the approach we suggest below because the assessment is based on predicted values from a model onto an equally spaced grid within the assessment region.

The area swept by the trawl is defined by the opening of the net (the wingspread) and the distance towed. Data on trawl distance, doorspread and wingspread of the net are not always available. (Wingspread can be approximated from doorspread when wingspread is not available). We have obtained this information in a number of ways. Specifically, distance is calculated from the location of the shoot and haul locations of the trawl. For GOV hauls, wingspread is calculated from the expression: wingspread = doorspread*0.1887 + 5.8228. Area swept is then approximated by distance * wingspread (O'Donoghue et al. 2018).

Figure 2 shows the different surveys which relate to the gear types in table 1. GOV is the dominant gear although other gears are used in the same North Sea region in which GOV is heavily used. For the assessment modelling, all predicted values need to be standardised to the same gear and the same area swept. The standardised gear should be GOV for the GNS, CS, and BB. The area swept is standardised to 66,814 m^2 for all regions.

3.5 Statistical methods

The important thing about any assessment method is that it should not produce biased results (see the three potential causes of bias outlined in Section 3.4). Because the same gear is not used in all assessment regions, the approach suggested for the 2021 OSPAR assessment for each year is as outlined below. Full details are in appendix 1 which refers to the methods from the OSPAR 2022 assessment (Barry and Russell et al., 2022).

1. Fit a logistic generalised additive model (with the response variable being presence (1) or absence (0) of the litter type being assessed) to map litter measurements over the regions using available data in any one year. This model includes area swept, latitude, longitude and gear type as explanatory variables.
2. Predict modelled values onto a square grid encompassing the assessment region. Currently this grid does not predict onto points that are further than 20 Km from an observed point.
3. The mean for the region is the mean of all the grid points within the region.
4. 95% confidence intervals are obtained using a procedure described in Wood (2017). This is essentially parametric bootstrapping of the model parameters.

3.6 Assessment measurements

For the latest assessment seven years of data was used as this was how much we had confidence in. A minimum of 6 years' worth of data could be used to match the MSFD period for future assessments. Although more data would be better to try and understand tends over time. All areas where there is enough data should be assessed and special maps used to show these trends for the OSPAR region.

The following groups are analysed in the 2022 OSPAR assessment: total count, Single Use Plastics (bags and bottles), fisheries items and the material categories (e.g., Plastic, rubber, metal, glass, and natural). These are listed in table 2. Work should continue between indicators to ensure that the categories used are comparable for interpretation of results. For example, the latest beach litter assessment (Lacroix, C. et al. 2022) used a maritime related item category instead of fishing due to difficulties in understanding the source of certain abundant items such as ropes which may come from fishing or shipping or another source.

Table 2: *Litter categories used in the latest OSPAR seafloor litter assessment (Barry and Russell et al., 2022)*

Variables used in the assessments	Items included
All litter	All items
Plastic	Bottles, sheets, bags, caps/lids, fishing line (monofilament and entangled), synthetic rope, fishing net, cable ties, strapping bands, crates and containers, diapers, sanitary towel/tampons and all other plastic items
Metal	Cans (food), cans (beverage), fishing related, drums, appliances, car parts, cables and all other metal items
Rubber	Boots, balloons, bobbins (fishing), tyres, gloves and all other rubber items
Glass	Jars, bottles, pieces and all other glass items
Natural	Wood (processed), natural rope, paper/ cardboard, pallets and other natural items
Fishing	Fishing line (monofilament and entangled), rubber bobbins, rope (natural and synthetic), fishing related metals, fishing net
Bags	Plastic bags
Bottles	Plastic bottles

Descriptive analysis of the items that occurred most frequently in all hauls in each of the three OSPAR regions, is used to provide the top 10 probabilities of litter items in these regions. This was used because historically, count data has not always been available and some of it was thought to be of poor quality. To get this data no attempt has been made to do any modelling of the data, to account for spatial bias or differences in haul characteristics (e.g., gear); these are simply what the raw data demonstrates. The implicit assumptions are that the probability that an item is detected is the same in different areas and is not affected by the gear type. In future years where there is confidence that the CEMP guidelines and ICES guidelines (2022) have been followed and the gear set up has remained constant over the assessment period then count can be considered to be used to show number of litter items. The current assessment has provided a case study for GNS for how this could be done.

4. Change Management

The Environmental Impacts of Human Activities Committee (EIHA) is responsible for policy and resource questions on the acquisition and storage of data. Quality control and analysis and technical aspects are managed by the Intersessional Correspondence Group on Marine Litter (ICG-ML) with the support of the Seafloor Litter Expert Group (SLEG). Any changes relating to data generation and management should be discussed and agreed with ICES Working Group on Marine Litter (WGML).

5. References

Barry, J., Russell, J., van Hal, R., van Loon, W.M.G.M., Norén, K., Kammann, U., Galgani, F., Gago, J., De Witte, B., Gerigny, O., Lopes, C., Pham, C. K., Garcia, S., Sousa, R., Rindorf, A. 2022. *Composition and Spatial Distribution of Litter on the Seafloor*. In: OSPAR, 2023: The 2023 Quality Status Report for the North-East Atlantic. OSPAR Commission, London.

Directive (EU) 2015/720 of the European Parliament and of the Council of 29 April 2015 amending Directive 94/62/EC as regards reducing the consumption of lightweight plastic carrier bags. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32015L0720.

Directive (EU) 2019/904 of the European Parliament and of the Council of 5 June 2019 on the reduction of the impact of certain plastic products on the environment. https://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:32019L0904&from=E.

Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32008L0056&from=EN.

Van Franeker, J. A., Kühn, S., Anker-Nilssen, T., Edwards, E. W., Gallien, F., Guse, N., ... & van Loon, W. M. (2021). New tools to evaluate plastic ingestion by northern fulmars applied to North Sea monitoring data 2002–2018. *Marine Pollution Bulletin*, *166*, 112246.

Galgani, F. *et al.* (2000) 'Litter on the sea floor along European coasts', *Marine Pollution Bulletin*. doi: 10.1016/S0025-326X(99)00234-9.

ICES (2018): Report of the Working Group on Marine Litter (WGML). ICES Expert Group reports (until 2018). Report. https://doi.org/10.17895/ices.pub.8243.

ICES. 2022. ICES manual for seafloor litter data collection and reporting from demersal trawl samples. ICES Techniques in Marine Environmental Sciences Vol.67. 16 pp. https://doi.org/10.17895/ices.pub.21435771.

ICES. 2022. Photograph guide for ICES manual for seafloor litter data collection and reporting from demersal trawl samples. In: ICES. 2022. ICES manual for seafloor litter data collection and reporting from demersal trawl samples. ICES Techniques in Marine Environmental Sciences Vol.67. 16 pp. https://doi.org//10.17895/ices.pub.21435771.

Lacroix, C., André, S., and van Loon, W. 2022. *Abundance, Composition and Trends of Beach Litter*. In: OSPAR, 2023: The 2023 Quality Status Report for the North-East Atlantic. OSPAR Commission, London.

O'Donoghue, A. & Van Hal, R. 2018. Seafloor Litter Monitoring. International Trawl Survey 2018. Wageningen, Wageningen Marine Research (University & Research centre), Wageningen Marine Research report C052/18. 58 pp.

OSPAR. 2021. Strategy of the OSPAR Comission for the Protection of the Marine Environment of the North-East Atlantic 2030. Agreement 2021-01: North-East Atlantic Environment Strategy (replacing Agreement 2010-03.

OSPAR. 2022. The second OSPAR Regional Action Plan on Marine Litter. OSPAR Publication 2022/891. [documents (ospar.org)]

Wood, S, N. (2017) Generalized Additive Models: An Introduction with R, second edition. CRC Press. 467.

Van Loon, W., Hanke, G., Fleet, D., Werner, S., Barry, J., Strand, J., Eriksson, J., Galgani, F., Gräwe, D., Schulz, M., Vlachogianni, T., Press, M., Blidberg, E. and Walvoort, D., 2020. A European Threshold Value and Assessment Method for Macro Litter on Coastlines. EUR 30347 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-21444-1, doi:10.2760/54369, JRC121707.

Appendix 1: OSPAR Assessment methods (technical supporting information for Barry and Russell et al., 2022)

The area swept by hauls clearly has an influence on the amount of litter collected. We want to be able to control for this variable in our assessments to make sure that any apparent differences in litter amounts are real effects and are not caused by differences in area swept.

We can calculate the area swept for a haul by the wingspan of the haul multiplied by the distance towed. For beam trawls, the wingspan is simply the width of the gear.

For GOV hauls, there are many situations in which wingspan is not recorded. Thus, for GOV hauls we used the formula provided by O'Donoghue and Van Hal (2018) to impute wingspan from doorspan for Dutch hauls:

$$Wingspan = 5.8728 + 0.1887 * Doorspan \qquad (1)$$

This formula is being used as an approximation for all GOV hauls in this assessment. Future assessments need to consider the ICES Swept Area Calculation Algorithms reports which describe mathematical functions for estimation of the missing values of doorspan, wingspan and distance for each country and for each survey. These mathematical functions are based on observed values over the years of the survey and are provided by the national institutes.

According to the Manual of the IBTS North-Eastern Atlantic Surveys Version 4.0 (ICES, 2017) the mean wingspan for NCT hauls (PT-IBTS) is 15.1. We thus used this figure for our data.

For our original data set, 48% of the 17,100 hauls had missing wingspan values. After the imputation procedures described above, this reduced to 5,856, or 34%, of the hauls. These hauls were not used in the analyses.

In terms of the gear characteristics of the UK NS-IBTS survey, we received the following advice from the operatives:

5. The net used is a GOV (Grand Ouverture Verticale) 36/47 Otter Trawl, where the 36/47 is the length of the top rope and the fishing rope, respectively.
6. The net is rigged to IBTS (International Bottom Trawl Survey) standard to allow for comparative catches across all IBTS surveys.
7. Focusing the dataset on 2015-2020 is advised, as significant changes were made to the gear just prior to this, relating to the net construction, as well as introducing a change to the floatation of the net.

With regard to the sample locations for the UK NS-IBTS survey:

- Surveyors always try to fish the same tow line as the previous year, if it was successful. However, this is not always possible due to issues such as tide. Where this happens, they always try to bisect the previous year's tow line.
- The survey requires only that a tow is in the correct ICES rectangle. Therefore, if the tow needs to be moved (due to previous gear damage, cables laid since the previous year, or an active seismic vessel etc.) it will be. This probably happened on less than 5 out of the 77 tows each year.

Statistical Methods

As explained, the challenge with the different gear types is that they all have different abilities to catch litter. Whilst we can control for the area swept of the gear, there will be other characteristics of the gears that will induce different litter catchability rates. The consequence of our data containing different gears is that we do not know if levels of litter amounts are a function of the haul location, or the gear type used. Having said that, we want to use all the available data, so that we can maximise our spatial coverage and intensity of sampling. We return to this issue below.

Another issue is that of unequal sampling effort in space. If one part of a region is sampled more heavily than another part, when we take a simple mean of all the results our mean will be biased towards the area with the greater sampling effort. Thus, for example, if we sample more heavily in an area with low litter levels than in parts of the region with higher litter levels, a simple mean will underestimate the mean litter levels per km^2 over the whole region.

For both the assessment of the probabilities of hauls containing litter items and the UK NS-IBTS demonstration study of litter counts, we adopted a statistical modelling approach to accommodate the three potential problems (area swept, gear type, and spatial sampling effort) described above.

Probabilities of hauls containing a litter item

The first step was to place a (virtual) square grid over the region. The reason for doing this was so that we could generate a grid of points evenly located over the region. For each year, we then established whether a grid point was less than 20 km from a sampling point. If this was the case, then that grid point was used as a point at which modelled values were estimated.

We fitted a Generalised Additive Model (Wood, 2017) to each year's data using the function *gam* in the *mgcv* R package. This model's link function was of the form:

$$f(p) = \alpha + \beta areaswept + gear_j + s(lat, long) \qquad (2)$$

where p is the probability of one or more litter items being present in a haul, $f(p)$ is the standard logit function for binary data such that $f(p) = \log\left(\frac{p}{1-p}\right)$, α is an intercept term, βareaswept is a linear function of haul area swept, $gear_j$ is an estimate of the jth gear effect, and $s(lat, long)$ is a smooth two-dimensional function of latitude and longitude that includes both the two main effects and the interaction. For the Bay of Biscay and Iberian Coast assessment only one gear was used – thus, for that assessment, the $gear_j$ term in (2) was not needed.

Parameters in eq. (2) for each year, 2012-2019, were estimated using the default method of cross-validation (Wood, 2017). Essentially, cross-validation works by leaving out each point in turn and calculating the mean ability of the remaining points to predicting the left-out datum. Parameter estimates are chosen that give the best mean prediction. We did not restrict the degrees of freedom of the smooth terms (the degrees of freedom determine how flexible – or wiggly – the smooth terms are).

Once models of the form eq. (2) had been fitted, we used the parameter estimates from the models to predict p, the probability that the grid point would contain a litter item. The latitude and longitude in the model were simply the location of the grid point, the area swept was set to a constant value of 57,000 m^2 for all grid points and the gear was set to the GOV gear if there was more than one gear in the region. The exception was for the small analysis conducted for the Iberian Peninsula. The NCT gear was used here, with a smaller area swept. Thus, for this analysis, the area swept was standardised to 44,000 m^2.

There is a caveat in terms of the area swept variable. For some analyses, area swept was not a statistically important variable in explaining litter proportions. For example, sometimes the parameter β was estimated as negative. We are not sure why this is the case; it could be that the area swept for hauls for those analyses were similar and it was difficult for the model to pick up any relationship between increasing area swept and increasing proportions of litter found. It is important to be wary of using the area swept variable for predictions in these situations. For example, if a negative β is used, this will impact on our predicted values of p. Thus, we adopted the approach below.

Once the model (2) had been fitted, we assessed the parameter β. If β was not statistically significant ($p < 0.05$) or if the estimate of $\beta < 0$ then the model was re-fitted, but without the area swept term. Prediction then used only latitude, longitude and, if relevant, gear type.

A similar issue arose if the spatial element of the model was not statistically significant. In such situations, showing differences in predicted litter levels over space is misleading given that we do not have evidence that such spatial differences exist. In these circumstances, we did not use the spatial term in our predictions. In situations where the area swept effect is not statistically significant and there is only one gear, this results in our predictions being the mean values of the litter (0,1) variable for the region – i.e. the variable has value 0 if a litter item is not found and value 1 if at least one litter item is found.

For mapping purposes, we simply used the predicted values at the grid points. We also calculated standard errors, that reflect the precision of these predictions, but these have not been used for the current assessments.

To get an estimate of p for a particular region and year, the mean of all predicted points on the grid was used. However, so that fair comparisons could be made between years, only grid points that were within 20 km of a sampled point in *all* eight years of the study were used (whereas, for the maps, all grid points for a particular year that were within 20 km of a sampled point were used). To get a 95 % confidence interval for this estimate we used the procedure described on pages 342-343 of Wood (2017). Essentially, this involves simulating from the posterior distribution of the parameters - assuming that they have a Multivariate Normal distribution, calculating the linear predictor (LP, the right-hand side of equation (2)) for each simulation and then transforming back to the original scale using the standard $p = \frac{\exp(LP)}{1+\exp(LP)}$ back transformation. For repeat simulations (we used 1000), the 95 % confidence interval is the 2.5th and 97.5th percentiles of these simulated ps. Essentially, we might think of this process as a form of parametric bootstrapping of the parameters.

Litter counts

We have not undertaken a full modelling assessment on counts following guidance from the working group. However, we have done a demonstration study for NS-IBTS surveys conducted in the Greater North Sea by UK. We have also done some preliminary, exploratory analyses of catchability of different litter types by gears. Clearly, these results should be interpreted with caution for the same reasons that we have not done a full counts assessment. However, we believe that they provide some interesting insights and so we report our initial findings in this assessment.

For the demonstration study of counts in the GNS, we modelled the total litter counts by a GAM model, with link function:

$$log(E[C]) = \alpha + \beta areaswept + s(lat, long) \quad (3),$$

where $E[C]$ is the mean count. We used the raw data to compare Negative Binomial and Poisson error distributions for the counts. Theoretically, we would expect the Negative Binomial distribution to be appropriate if litter items were clustered on the sea floor; the Poisson distribution would be appropriate if litter items were spatially random.

For the study of catchability of different litter types, we transformed the original counts to count per unit effort by dividing the count by the area swept (in m^2) and then multiplying the answer by 10^6. We thus have counts per 1km square.

To calculate the conversion factor between Beam Trawl (BT) and GOV hauls for the GNS, eighty spatial squares covering the GNS region were chosen. The mean of the ratio (MOR) of BT to GOV counts per unit effort was calculated over these 80 squares for each of the litter types defined (this was done for all data points and was not done separately for each year). Thus, to standardise counts per unit effort to GOV hauls, all BT counts per unit effort values were divided by the appropriate MOR. For the Celtic Seas, 90 squares were used, and ratios calculated between BT and GOV, and between PORB and GOV. For the Iberian Coast and Bay of Biscay, 80 squares were used, and ratios calculated between BAK and GOV gears.