Environ Monit Assess (2020) 192:648
https://doi.org/10.1007/s10661-020-08602-w

An open-source geospatial framework for beach litter monitoring

Jessica L. Schattschneider · Nicholas W. Daudt · Mariana P. S. Mattos · Jarbas Bonetti · Nelson Rangel-Buitrago

Received: 16 March 2020 / Accepted: 10 September 2020

Abstract Here, we present a framework for a beach litter monitoring process, based on free and open-source software (FOSS), which allows customization for any sampling design. The framework was developed by means of a GIS project (QGIS), a GIS collector (QField), and an R code, allowing further adjustments according to the area to be surveyed and research questions. The aim is to improve data collection, accessibility, and interoperability, as well as to help to fill the currently existing gap between fieldwork and data analysis, preventing typos and allowing better data processing. Therefore, it is expected to take less than an hour from ending fieldwork to obtaining up-to-date products. To test the developed open-source geospatial framework, it was applied in different sectors and dates on an important southern Brazilian touristic beach. Results obtained from the open-source geospatial framework application produce baseline information on beach litter issues, such as amounts, sources, and spatial and temporal patterns. Adoption of the framework can facilitate data collection by local and regional stakeholders, and the results obtained from it can be applied to support management strategies. For researchers, it produces spatialized data for each item in an already tidy format, which can be used for robust and complex models. A series of supplementary files support reproducibility and provide a guide to future users.

Keywords Environmental monitoring · FOSS · GIS · Marine debris · Plastic · Polluting sources

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10661-020-08602-w) contains supplementary material, which is available to authorized users.

J. L. Schattschneider · J. Bonetti
Laboratório de Oceanografia Costeira, Centro de Ciências Físicas e Matemáticas, Universidade Federal de Santa Catarina (UFSC), Florianópolis, SC, Brazil

J. L. Schattschneider (✉) · M. P. S. Mattos · J. Bonetti
Programa de Pós-Graduação em Oceanografia, Centro de Ciências Físicas e Matemáticas, Universidade Federal de Santa Catarina (UFSC), Florianópolis, SC, Brazil
e-mail: jessica.leiria@gmail.com

N. W. Daudt
Programa de Pós-Graduação em Oceanografia Biológica, Instituto de Oceanografia, Universidade Federal do Rio Grande (FURG), Rio Grande, RS, Brazil

N. W. Daudt
Museu de Ciências Naturais (MUCIN), Universidade Federal do Rio Grande do Sul (UFRGS), Imbé, RS, Brazil

M. P. S. Mattos
Laboratório de Gestão Costeira Integrada, Centro de Ciências Físicas e Matemáticas, Universidade Federal de Santa Catarina (UFSC), Florianópolis, SC, Brazil

N. Rangel-Buitrago
Departamentos de Física y Biología, Facultad de Ciencias Básicas, Universidad del Atlántico, Barranquilla, Atlántico, Colombia

Introduction

Beach litter monitoring schemes are based on surveys to count and classify items, and draw profiles concerning the most critical accumulation points, and types of items

Published online: 19 September 2020

and materials (Moreira et al. 2016; Schulz et al. 2019). Monitoring is the basis for the determination of amounts, distribution, polluting sources, and trends of litter in the marine environment, and provides the baseline for comprehension and further solutions for beach litter issues (Ryan et al. 2009; Schneider et al. 2018; Schulz et al. 2019). With the significant increase of beach litter reports (Derraik 2002; Schneider et al. 2018), new technologies seeking efficiency in time and cost are being adopted in monitoring efforts, increasing the reliability and replicability of results, and decreasing human bias (e.g., Martin et al. 2018). Despite the increase of new methodologies, beach monitoring based on human observation is still the most popular approach for beach litter evaluations.

Geographic Information System (GIS) platforms bring unique opportunities to implement beach litter monitoring, retrieving clues for determining litter spatial distribution and polluting sources. Besides that, the adaptability of GIS-based frameworks allows the minimization of data requirements and they perform as comprehensive and widely accessible planning tools (Haarr et al. 2019). Mobile devices now have devoted applications for litter monitoring information (e.g., Marine Debris Tracker, CleanSwell, Marine LitterWatch). These applications integrate the data in a GIS so that one can collect beach litter data from anywhere with mobile Global Navigation Satellite System (GNSS) and download it to an online repository. The use of GIS is being proposed to map marine debris and increase the participation of local people to monitor and understand the issue (Bennett-Martin et al. 2016). Moreover, it supports beach clean-ups by analyzing predictable areas of litter accumulation (Haarr et al. 2019). In addition to increasing efforts on beach litter surveys, these approaches can be powerful tools to engage and sensitize the public and to generate alerts for the environmental concerns of litter impacts. However, being based on proprietary software, these platforms do not present a flexible structure, preventing the adaptation of monitoring design for specific purposes or adjustments towards the improvement of predefined analysis.

Currently, the existence of several independent monitoring methodologies is another noticeable difficulty in mapping global distribution and trends of litter along with the beach environment (Galgani et al. 2015). Given the deficiency of minimal data and structure—or even the absence of data—the integration interoperability and reproducibility of these efforts are a barrier to the worldwide understanding of beach litter patterns. In accordance, the lack of data accessibility and the absence of a transparent method are pointed out as key factors for the ongoing reproducible crisis observed in a variety of research fields (Baker 2016). In an attempt to change this scenario, the adoption of Findability, Accessibility, Interoperability, and Reusability (FAIR) guiding principles for research data stewardship has become a priority for the development of open science (Wilkinson et al. 2016). Beyond several benefits from FAIRness research, scientists can avoid spending a long time trying to integrate data from different datasets, favoring their focus on data analysis. In monitoring efforts in the marine environment, policies to implement these principles have been widely encouraged and increasingly required by many organizations; however, their implementation is still a challenge (Tanhua et al. 2019). This weakness is clear: a set of 35 marine research datasets frequently used for marine litter have all failed on evaluation of their level of FAIRness (Garcia-Silva et al. 2019).

In order to make FAIR principles applicable to reality, this concept should be applied broadly to all objects that are essential to the practice of research, including software and workflows (Collins et al. 2018). The adoption of free and open-source software (FOSS) can play a key role in the documentation of beach litter datasets, supporting flexible and transparent methodologies. The improvement of knowledge sharing is another advantage of using of these facilities, as is the move towards research reproducibility.

In this sense, this work presents an open-source geospatial framework for beach litter assessment and monitoring based on items' composition and their related polluting activities (as sources), using only FOSS facilities in all parts of the process. The framework is based on a GIS project, a GIS collector, and an R code, allowing further adjustments according to the area to be surveyed, the sampling design, and research questions. The aim is to improve data collection, accessibility, and interoperability, also helping to fill the currently existing gap between fieldwork and data analysis, preventing typos and allowing better data processing. We evaluated the usefulness of the open-source geospatial framework by applying it to an important southern Brazilian touristic beach. Results obtained from application of this framework provide comprehensive basic information and spatiotemporal beach litter patterns.

Materials and methods

Framework structure overview

The set of software that this framework integrates is based on its free use and transparency due to open-source codes, characterizing it as FOSS. This allows reproducibility and plasticity of any type of beach litter monitoring. Here, QGIS 2.18.7 (QGIS Development Team 2018; www.qgis.org) and QField v. 1.0.0 rc–5 (www.qfield.org) were used as geospatial tools, and data were wrangled (cleaned and transformed; Whickham and Grolemund 2017) and analyzed in R environment v. 3.5.3 (R Core Team 2019; www.r-project.org) due to their interoperability.

Figure 1 presents the framework structure developed in this paper, with three parts. The first part is related to the sampling design and creation of the GIS project, where one defines the corresponding files and attributes to be filled in the project, which considers the characteristics of the study area and the objectives of monitoring beach litter (Fig. 1a; Table 1). In the second part, the GIS project is uploaded to a GIS collector mobile app to be used in beach litter monitoring fieldwork (Figs. 1b and 2). In the third part, all data collected are directly imported to a PC and read through an R code, generating the result outputs (Figs. 1c, 3, and 4). A set of supplementary files and the R code comments are the supplements to this paper (Online Resources 1–3), which aims to support the use and implementation of this reference framework in any area.

Below, we detail the specific parts of the framework and how we implemented it. For a similar scenario and sampling design to our study case (see the "Applying the framework: a study case in southern Brazil" section), the framework is ready to use with products available in the Online Resources (i.e., the GIS project in Online Resource 2, and R code in Online Resource 3).

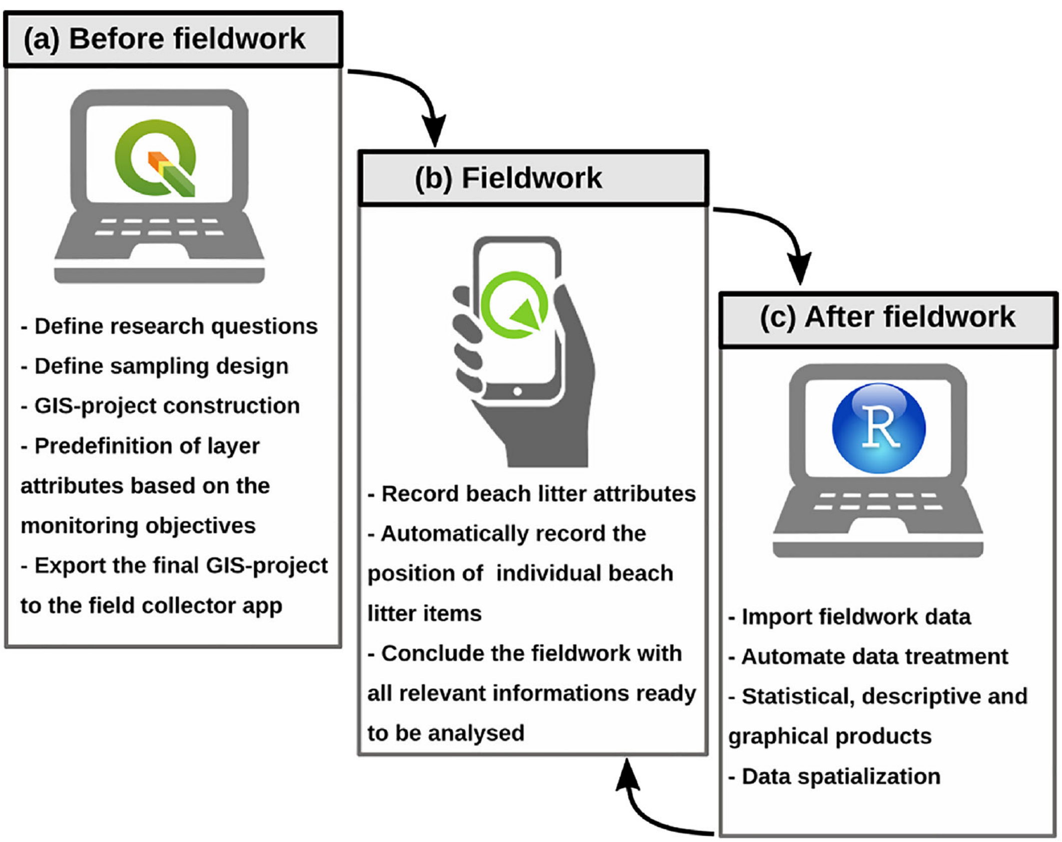

Fig. 1 Open-source geospatial framework for beach litter monitoring, based on free and open-source software. The framework is structured into three parts: **a** before fieldwork—define research questions, sampling design, and build a GIS project (in QGIS); **b** fieldwork—fieldwork and data acquisition (with mobile app QField); and **c** after fieldwork—data wrangling and analysis (under R environment)

Before fieldwork: sampling design and GIS project

An optimal sampling protocol is based on clear research questions, which will provide a guide to build up an effective GIS project. Here, the GIS project considered fundamental questions for a variety of beach litter monitoring schemes (Table 1) and three sampling sites, and focused on macrolitter pollution and its related polluting activities (see the "Applying the framework: a study case in southern Brazil" section, and the complete study case in Online Resource 4). The framework is open and suitable for different protocols, and researchers/stakeholders can modify it in accordance with their situation. Below, we point out some main subjects that must be taken into account for the definition of sampling design if one needs to adapt to the structure of the GIS project.

Table 1 Research questions that underpin baseline information on beach litter monitoring schemes, and how to achieve and visualize the answers

Question	How to achieve	How to visualize
(1). Which is the most littered site?	Count beach litter in different sites.	Summary table
(2). Which is the most common item?	Use an item-list base and count each item. Preferably, use an established list.	Summary table
(3). Which polluting activity contributes most to beach litter?	Define "polluting activity" sources for each item (e.g., dumping, smoking-related).	Histogram and ANOVA
(4). Is there a temporal pattern related to litter amounts, and the proportion of polluting activity inputs on each site?	Count beach litter in different periods of the year (months, seasons, etc.).	Trend lines graph
(5). Are there litter accumulations in specific beach zones?	Count beach litter in different beach zones (e.g., close to sand dunes vs. close to swash zone).	Dot graph (mean ± s.d.) and *t* test
(6). Do items from specific polluting activities accumulate in specific beach zones?	Count beach litter in different beach zones (e.g., close to sand dunes vs. close to swash zone).	Treemap plot
(7). How is the spatial distribution of the individual items related to their polluting activities?	Record each item individually, geospatialized.	Spatial representation

The number of sampling units can differ along with areas. For example, coastlines that are sheltered, with small beaches (2–5 km), can be sampled with just one transect (e.g., Corraini et al. 2018). On the other hand, medium (5–10 km) to large (> 10 km) exposed coastlines may need several transects for precise representation, according to the scale of the work (e.g., Araújo et al. 2018). Other aspects are the beach width and the local tidal regime because both influence litter accumulation and distribution processes. Therefore, it is necessary to consider the morphology and dynamics of the beach to optimize the monitoring protocol (Araújo et al. 2006; Velander and Mocogni 1999). Likewise, accumulation of litter by natural (e.g., currents) or anthropogenic factors (e.g., beach users) must also be taken into account (Santos et al. 2005; Velander and Mocogni 1999).

Other factors to be defined in the design of the assessment are, naturally, the size and typology of litter to be recorded. Parameters like size class (macro-, meso-, and microlitter, e.g., according to Ryan et al. 2009) and the type of materials to be collected (only plastic or all litter types) must be defined before fieldwork. These parameters will influence the definition of the litter list to be used in the field—those chosen need to agree with local reality and the research questions. For instance, the focus could be the major source of the litter (e.g., Araújo et al. 2018), the litter type (e.g., primary material, Topçu et al. 2013), or litter categories (e.g., individual items, Schulz et al. 2019).

All these aspects need to be carefully defined because wrong selections will generate additional time requirements and additional logistics to be implemented during sampling, as well as the loss of data quality.

Once the sampling design is defined, the QGIS project can be built with layers and respective attributes (Online Resource 1). These attributes include information about the litter items recorded (item, number, pictures), and their relation to the sampled location (site, date, transect). In the GIS environment, general rules can be defined for specific attributes related to the data type (e.g., integer, character, list of values, images) and restrictions at the moment of filling in the form in the field (e.g., number of characters, and if mandatory). Afterwards, the project is imported to a GIS collector.

Fieldwork: data collection

The geospatial collector mobile app QField is the GIS collector used for fieldwork, excluding the need for specific spatial instruments (e.g., a GNSS receiver), and ensures data standardization. QField can be freely downloaded in Android app stores and allows importation of the GIS project characteristics defined in QGIS. In the field, the user simply needs to run the application and fill in the form (Fig. 2).

A step-by-step protocol for GIS project construction and QField use is available in Online Resource 1.

After fieldwork: data analysis

After fieldwork, the collected data are imported from the GIS collector and processed on a PC under an R environment, a FOSS for statistical and graphical analysis. The use of GIS tools allows the collection of a considerable amount of different data, and one of the main advantages is the possibility of regularly updating the initial information with new sets of data (Nunes et al. 2009). Accordingly, we propose an R code which automates the process of importing new data, updating the dataset, and generating up-to-date analyses. The code is structured in sections related to different objectives (Fig. 3, Online Resource 3), which are detailed next.

Section “00” always needs to be run and is where the auxiliary tables are defined, i.e., what the surveyed beaches/sites and beach zones/areas are called, and the link between the litter list and its related polluting activities as specified in the sample design. Section “01” needs to be run just once, to upload data from the first fieldwork survey. Section “02” is for uploading the latest and previous dataset, and has to be run after each new fieldwork survey. Section “03” is where the latest data and auxiliary tables are merged, as well as the update of the dataset. Section “04” is for the analysis and outputs.

The products of Section “04” refer to the table, graphical, and statistical analysis produced based on the research questions (Table 1). A summary of the visual products generated by the framework to respond to these questions is presented in Fig. 4, where colors represent different polluting activities, capital letters different sites, and lowercase letters different areas. The table returns summaries for each item per site, which helps to visualize the most recurrent litter item and the site with most litter (Fig. 4, “01” and “02”). Concerning the graphs obtained, the first is a histogram that represents the relative contribution of different polluting activities to the amount of litter recorded for each site (Fig. 4, “03”). The second graph is a line plot that represents the temporal trend of the total litter items, and the contributing percentage of specific polluting

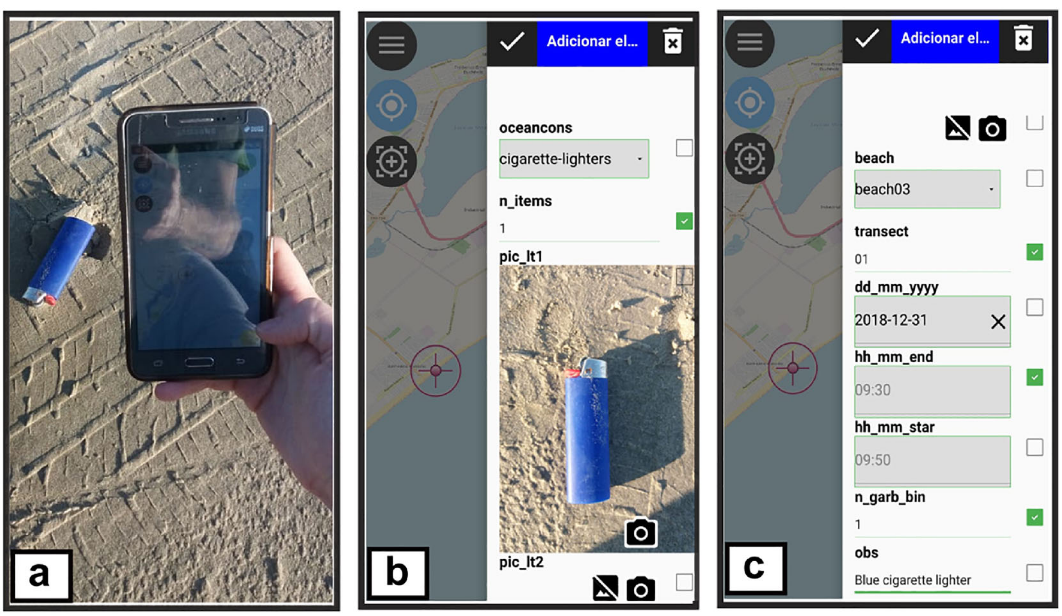

Fig. 2 Mobile app QField in fieldwork. **a** Record the item. **b**, **c** Fill in the attributes

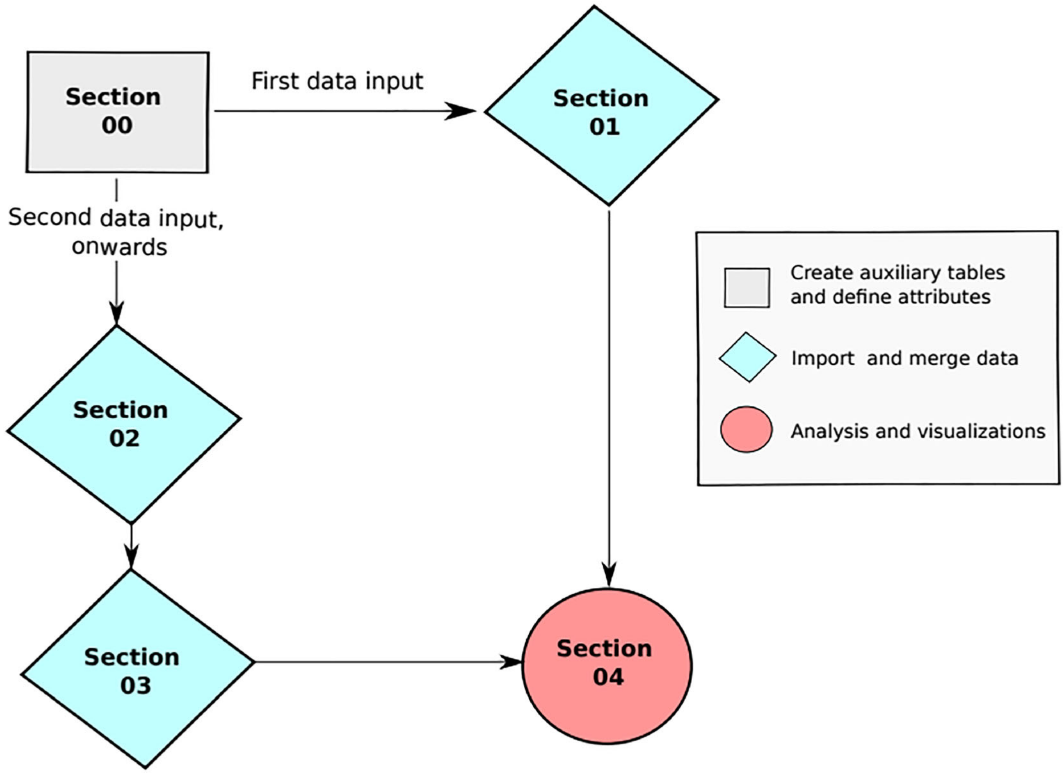

Fig. 3 Code structure and section flow. The code contains four different sections, and the sequence to be followed depends on the data input moment (first data input or second data input, onwards). Different symbols/colors are related to major objectives: square/gray represent the section where auxiliary tables are created and attributes are defined; diamonds/light blue represent the sections where data are imported, updated, and merged with auxiliary tables; and circle/red is related to the data analysis and visualization section

activities in the surveyed sites (Fig. 4, "04"). The third and fourth graphs are related to different zones (see the "Applying the framework: a study case in southern Brazil" section); the former is a dot graph with the mean ± standard deviation (s.d.) of the total number of litter items in each zone per beach (Fig. 4, "05"). The latter is a treemap plot that highlights the most representative polluting activity recorded in the surveyed site, biased by the total values obtained from both zones and within each zone (Fig. 4, "06"). The final visual product is a spatial representation of litter by means of bubbles, with size related to the number of items, and colors linked to specific polluting activities (Fig. 4, "07"). This map is also available in an HTML format, which can be accessed through any device connected to the internet, though it can be exported as a simple PNG format.

The reproducible R code is in a supplementary file (Online Resource 3), and specific R packages with their versions are cited along with the code. We have included an additional section ("Reproducible") that provides raw data from our trial, so one can reproduce the results presented in Online Resource 4—Section "Reproducible" should be run after Section "00," and followed by Sections "03" and "04."

Applying the framework: a study case in southern Brazil

Applicability of the framework was tested in a 23-km coastal segment located at Praia do Cassino (c. 32° 11′ S; 52° 10′ W), on the longest uninterrupted sandy beach in Brazil, 220 km long. This subtropical beach is a wave-dominated environment which experiences a micro-tidal regime and whose water levels are strongly influenced by storm winds from the south (Seeliger et al. 1997). The tourism industry represents one of the most important activities of this area, and during summer, the local population can increase significantly. At Praia do Cassino, cars are allowed to transit the beach year-round, which coupled to the beach length leads to uneven patterns of human-related occupation/activities along the

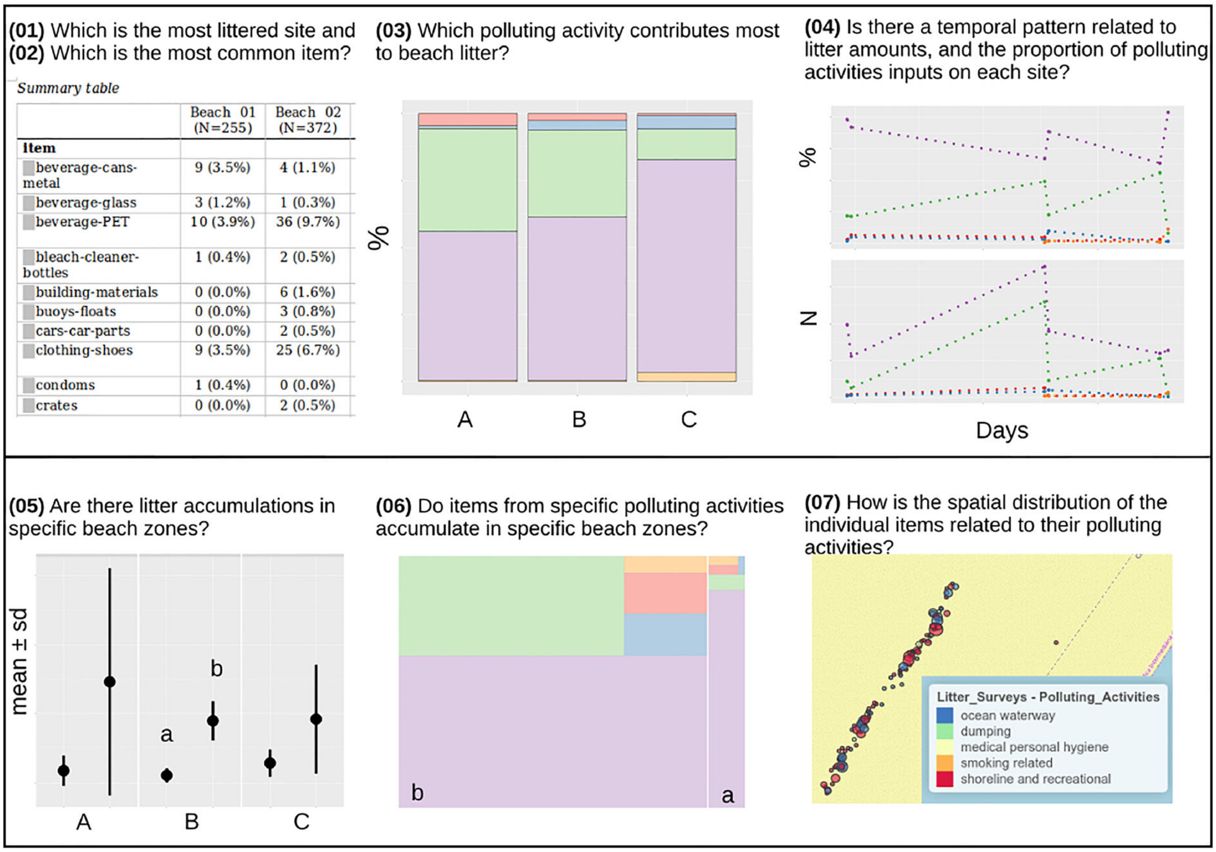

Fig. 4 Summary of graphical and descriptive results based on the open-source geospatial framework to achieve the research questions proposed in this work, encompassing spatiotemporal attributes. Numbered questions are the same as those in Table 1. Colors are related to different polluting activities (in questions 03, 04, and 06: orange—smoking-related, purple—shoreline and recreational, green—ocean waterway, blue—dumping, and red—medical and personal hygiene; question 07 has its own color legend), capital letters (A, B, C) represent different sites, and lowercase letters (a, b) represent different beach zones. This is a generalized figure to exemplify the framework products; the specific results from this work are in Online Resource 4

beach. The distribution of beachgoers is highly concentrated near the main road, and going south, visitor density decreases significantly. Detailed information on the study area can be found in Online Resource 4.

Following the steps of each part of the proposed framework, we first defined the baseline questions (shown in Table 1) and the sampling design, which took into account the study area characteristics. Briefly, three transects were carried out in three sites ~ 7 km apart; each transect had a sampling area of 400 m^2 (totaling 1200 m^2/site, covering 3600 m^2/survey), with two transects on the upper beach zone (close to the foredunes) and one transect on the lower beach zone (~ 3 m above the swash zone) (Fig. 5). Macrolitter (> 10 cm) was chosen as the target, and the International Coastal Cleanup–Ocean Conservancy (ICC/OC) list (Ocean Conservancy 2010) was used as the basis for individual litter types and their related polluting activities (Table 2). To encompass a temporal attribute, we conducted three surveys from early to late austral summer (i.e., December, February, and March) 2018/2019.

This information was the base for the GIS project (available in Online Resource 2) following the step-by-step protocol available in the supplementary file (Online Resource 1). The GIS project was designed with only one editable layer, including 12 attributes related to the sampling location (beach, transect, date, and hour), litter information (type, amount, pictures), number of garbage bins available in the transect, and additional observations. Closed forms were used to record attributes related to the sampling site and to the litter items; a pop-up calendar was used to fill in the date attribute. Also, we included an uneditable OpenStreetMap layer basemap in the project, improving the user interface. For the litter

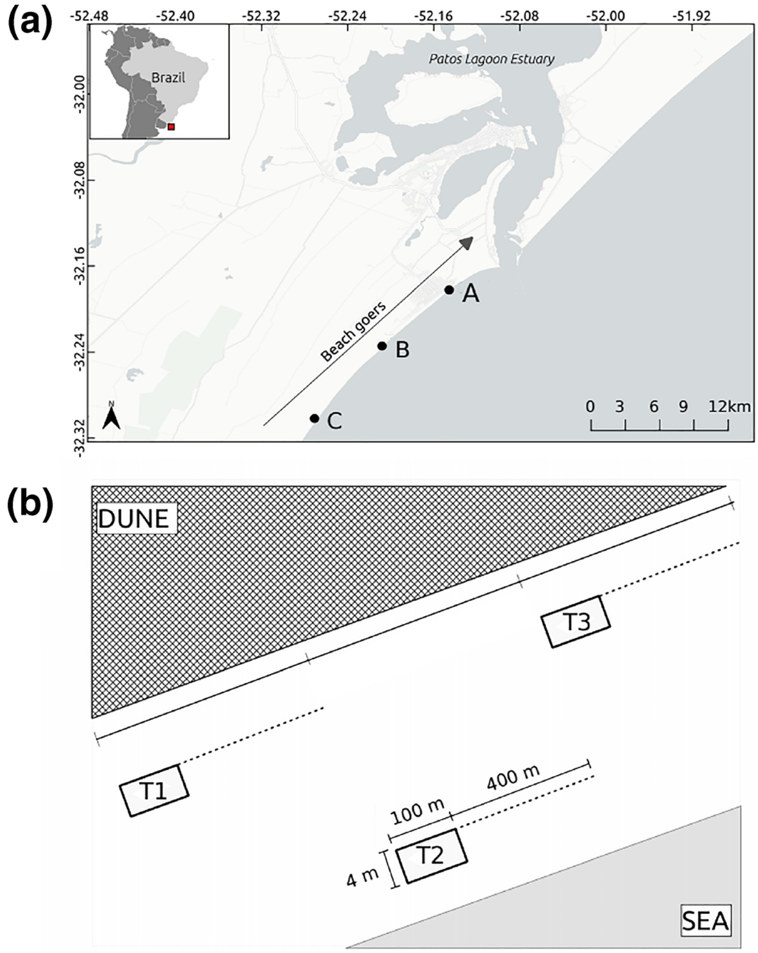

Fig. 5 Sampling design applied for litter quantification in the study case at Praia do Cassino, southern Brazil, based on the open-source geospatial framework. **a** Study area showing the location of the three sampled sites (A, B, and C), with arrow indicating beachgoers concentration. Each site was sampled according to the scheme presented in **b** where the rectangles represent the three transects (T1, T2, and T3) with areas 100 m long × 4 m wide, 400 m apart. T1 and T3 are located in the upper beach zone, and T2 in the lower beach zone

Table 2 Categories of polluting activities implemented in the framework structure, illustrating their main sources (based on Ocean Conservancy 2010)

Activity	Description
Shoreline and recreational	Mainly littering by beachgoers, but also from waterside sports and that carried from streets, drains, and gutters.
Ocean/waterway	Solid waste from recreational fishing/boating, recreational/commercial fishing and shipping, and from the oil and gas industry.
Smoking-related	Smoking-related packaging and materials.
Dumping	Building and construction materials, tires, cars parts, drums, household trash, and appliances.
Medical and personal hygiene	Materials discarded into the sewer system, toilets, or left behind by beachgoers.

list, we included the main polluting activities as an option, to ensure that all items could be recorded without the need to create new item classes.

Accordingly, this project was then imported to QField, used to record all macrolitter in each transect (Fig. 2). By simply noting and checking the cells of the GIS project in QField, beach litter data were recorded, and the output was tidy and ready to be wrangled. Then, using the R code at the end of each fieldwork period, we kept the dataset updated and reproduced the results proposed (i.e., Fig. 4). These six products respond to basic answers related to the spatiotemporal pattern and qualification of the beach litter.

Results

The framework allowed easy and simple recording of beach litter data and produced fast and comprehensive results. The main products generated by the open-source geospatial framework for beach litter monitoring are presented in a general perspective in Fig. 4—colors representing different polluting activities, capital letters different sites, and lowercase letters different zones.

No problem was observed in any of the three parts of the framework, and all our research questions were satisfied. The GIS collector QField had no crashes when recording the litter items during fieldwork. The only unexpected crashes occurred in a few cases during the first fieldwork survey, with pictures not being associated with the recorded items. The workaround to fix this issue was the addition of two new attributes to the GIS project (i.e., "pic_lit2," "pic_lit3"). In the third part related to data processing (Fig. 1c), the auxiliary tables were created based on the litter classification list, and the data from each new fieldwork survey were merged with the previous ones without problems. This process permitted up-to-date analysis including the whole dataset after the addition of any new fieldwork data. From the end of data collection (fieldwork) to producing the final products of the up-to-date results, we took less than an hour.

Results obtained from the framework application provided spatiotemporal beach litter patterns. As a summary of our study case, 855 macrolitter items were recorded from the 10,800 m^2 of sandy beaches, considering all surveys. Beach litter was encountered in all sites and during every sampling period. Thirty different litter categories (types) were observed, the most abundant being plastic-related items. Plastic bags, food packaging items, fishing lines, ropes, straw stirrers, and beverage PET dominated the samples, representing almost 70% of all recorded items. Differences were observed between the three monitored sites in Praia do Cassino, with the northern and southern sites being more littered than the central one. While the northern site had the majority of items related to shoreline and recreational activities, the southern one had a notable influence of litter from ocean and waterway activities. Temporally, the graphs were suitable for tracking differences in the amount of litter. Differences in accumulated amounts between zones were evident, though the proportional composition of related litter polluting activities was basically the same. The full results of our trial, with individual graphs and products and a brief discussion focused on the local relevance, are available in Online Resource 4.

Discussion

Framework evaluation

The framework proved to be a simple and useful tool to collect and analyze data and provide insights into spatiotemporal patterns in beach litter, being successfully applied at Praia do Cassino. The use of FOSS benefited the framework's interoperability, improving the data flow from collection to the final products. Furthermore, the absence of typos, coupled with the structured dataset returned by the GIS collector, shortened the time between fieldwork and data analysis.

The use of geospatial facilities to map environmental issues is increasingly being adopted, such as for marine pollution (Bennett-Martin et al. 2016; Cowger et al. 2019; Haarr et al. 2019). For beach litter monitoring, previous efforts created new tools for litter collection (e.g., Jambeck and Johnsen 2015) and analysis (e.g., Schulz et al. 2019). Although we have not created a new tool, to the best of our knowledge, this is the first attempt to integrate existing and widely used open-source software in a workflow focused on beach litter monitoring. The proposed framework comes as an effort to improve FAIRness in the subject of marine litter (Garcia-Silva et al. 2019), by implementing "Interoperability" and "Reusability," as well as reproducibility of the whole process. On the other hand, adding a synchronized database could benefit the "Findability" and "Accessibility" components of FAIR principles for these efforts (Wilkinson et al. 2016).

Geospatial approaches are crucial to prevent loss of data needed for beach litter monitoring, improving understanding of the problem and searching for solutions (Jambeck and Johnsen 2015). In this sense, the efficiency in data collecting time and the set of final products makes the framework proposed in this work a potential tool to be used in monitoring schemes. Spreadsheets are still the most used tool for the storage of data obtained in beach litter monitoring. However, their use requires the consideration of organizing premises set to avoid input data errors (Broman and Woo 2018). Common issues in traditional monitoring processes (i.e., paper-based protocols stored in spreadsheets) can be avoided by the adoption and use of geospatial tools (Gouveia et al. 2004). In this context, we had a substantial decrease in data loss due to typing errors (in fact, we had none), which is commonly observed when spreadsheets are used. Furthermore, the use of a GIS collector noticeably reduced the timespan from fieldwork data acquisition to final products when compared to the traditional litter monitoring process.

Another characteristic of the traditional approach for beach litter monitoring is the association of all data collected to a single coordinate related to the general location of the beach sampled, losing the spatial information of individually collected items. However, at a local scale, litter monitoring based on GIS collectors can return spatialization for individual items, which can help to identify points of accumulation (e.g., improving input data in predictive models; Haarr et al. 2019). Gathering continuous monitoring data with a GIS collector such as QField, one can also precisely infer removal and lateral transport rates of beach litter. Additionally, geotagged images can be associated with individual records, facilitating further analysis—and re-analysis. For instance, an image associated with the litter record can easily be reclassified to incorporate another database for further/other analysis (e.g., Martin et al. 2018).

Consideration of the reduction of eventual input errors and the use of common standards is necessary when adapting this open-source framework to different realities and scenarios. The use of recognized litter classifications (e.g., OSPAR list, ICC/OC list) should also be considered to keep a general standard that enables eventual comparison of results with other surveys, and for meta-analysis. The litter categories used in this work were adapted from the ICC/OC list which is composed of 47 classes. In this classification, some locally common items mostly related to recreational (e.g., coal for barbecues) and religious (e.g., candles and religious offerings) activities could not be recorded as no existing category satisfied the item type. For instance, in these cases, the general classes of "polluting activities" were considered, attempting to preserve polluting source information and maintaining the original list (e.g., instead of creating a new field for "coal," we recorded it as its general polluting activity source, i.e., "shoreline and recreational" activity). However, independent of the list of categories, we indicate closed forms in any situation as it minimizes eventual typing errors and avoids confusion during the data analysis process.

Some considerations must be made regarding the framework's cons. QField lacks a tracking tool to measure the distance monitored, implying the need for a complementary facility for this task or the use of fixed reference points. As alternatives, one can use the measurement tool to define the final point during fieldwork or create an additional layer (e.g., a polygon layer) containing the location of predefined monitoring sites. Additionally, unexpected crashes can occur with loss of pictures that were not stored (saved) with their specific picture attribute. This required the addition of extra fields of this type in the attribute table (e.g., as explained in the "Results" section, "pic_lit2" and "pic_lit3," instead of just "pic_lit1," in Online Resource 2), solving the problem. Besides that, the layout defined to represent variations in the amount of litter per time period among different sites has to be used carefully (i.e., Figure S4.7 in Online Resource 4). To avoid incorrect graphical products, we point to the necessity of finding a balance between time series resolution and the number of sites considered in the visualization. If a more extensive temporal analysis is to be produced, for a broad set of sites, some small code adjustments can quickly satisfy this demand. As already stated, a synchronized database could benefit data transfer from the collection device(s) to the PC on which the analysis will be run, speeding up the analytical part even more.

The decision to use the QField collector was made due to its compatibility with QGIS, although it is still not compatible with Apple devices. Besides that, we consider that the GIS collector version used here presented minimal problems even though it was in an experimental version. A stable version is now available, with fixed bugs and improved user interface. Based on our study case, the choice of a clean interface using only one editable layer made the recording process straightforward, simplifying it and making it intuitive for users with basic GIS experience. On the other hand, although the R interface is based on coding, it is a powerful tool for automating data wrangling and visualization, and can embrace complex data analysis for larger datasets. It also handles a variety of data formats, including the output data generated by QField, fitting the purpose of the presented framework.

Potential framework applicability

The presented framework highlights some fundamental questions related to the beach litter issue that can be applied to a variety of objectives, such as research, management, and engagement of the public through citizen science projects. The answers to these basic questions (Table 1) can support the development of management strategies. According to Williams and Rangel-Buitrago (2019), adequate litter management begins with accurate characterization of the magnitude, extent, sources, and impacts of litter (detailed knowledge of the problem). Once the typology and dynamics of litter are known, stakeholders may correctly intervene in the management practices to be developed. For example, identifying litter hotspots is one of the starting points for stakeholders (Rangel-Buitrago et al. 2017). Besides that, a lack of information (mainly on the temporal scale) prevents the creation and evaluation of local/regional policies (Hastings and Potts 2013). In this sense, the proposed framework, based on FOSS facilities which imply no costs, can provide the basis for management. Therefore, stakeholders from small regions and municipalities—which usually do not have a budget for beach litter monitoring—can make use of the proposed framework to obtain basic information on this issue, and address strategies to cope with beach litter. In fact, everyone who wants to find out answers to basic questions on beach litter issues in "three sites," with "three samples in each site," and using the "ICC/OC items-list," has access to the proposed GIS project and R code in a ready-to-use format (see Online Resources 2 and 3). As it is, the GIS project and the code should help gather baseline information at the local/regional level and can support management strategies at small to medium scales.

Projects involving local people as volunteers can also be favored. Citizen science is a way to engage and inform the public with science and can lead to reliable results in a variety of environmental issues, coupled with educational outcomes (Dickinson et al. 2012). Following the trend, beach litter data have increasingly been collected by the commitment of local people (e.g., Bennett-Martin et al. 2016; Cowger et al. 2019; Zettler et al. 2017). The use of mobile platforms engages the public, and nowadays, smartphones work as "networked devices" for data collection (Newman et al. 2012). The mobile app used in the framework can facilitate public participation in beach litter collection while maintaining standardization of the data, which in turn enables fast updates of the dataset and results. The georeferenced point of each item is also preserved by the mobile app, which can be advantageous for research questions, and to map specific items or polluting activities. The results can be updated even at the beach, and the public can be aware of the results they have achieved in less than an hour. Mobile platforms can be a powerful tool to raise awareness and to expand the litter issue to beachgoers. Moreover, citizen science can help to underpin data for management strategies (Hyder et al. 2015).

For research purposes, the advantages of these methods are noticeable. The use of FOSS promotes transparent collection, processing, and analysis of data, going towards FAIR principles (e.g., Collins et al. 2018; Wilkinson et al. 2016). Furthermore, FOSS has the inherent malleability of open-source software, providing the advantage of transparent adaptability in each step of the framework. These intrinsic characteristics can help researchers succeed in the metadata transparency so needed for future meta-analysis, aiming to understand the litter issue worldwide (Galgani et al. 2015; Ryan et al. 2009). Additionally, geotagged images can be processed by artificial intelligence algorithms, like machine learning (e.g., Martin et al. 2018), bringing another opportunity to re-use and re-analyze the data. Data collected through the framework give a tidy dataset, ready to use in data wrangling and interpretation.

Here, we propose an R code for data processing, given its transparency through its open codes. In the same way, Schulz et al. (2019) implemented an R package to analyze data from OSPAR beach litter monitoring, covering some basic questions and statistics on the issue. Our work goes towards the efforts of these later authors, and with minor adaptations, the data collected through the framework can be processed by their package (Schulz et al. 2019). Thus, with the addition of a polluting activity variable, OSPAR data can also be analyzed by our proposed code, due to the existing data interoperability.

Additionally, the dataset updates and fast results are significant framework improvements in data processes and analytics over traditional beach litter monitoring schemes. Here, we proposed a simple dataset update procedure and ran comprehensive baseline graphs and tables in the code, focused on polluting activities, bearing in mind management purposes and different stakeholders as users. Researchers can easily modify the code for more robust analyses (e.g., regression models, time-series analyses, numerical modeling). The nature of the framework is open, and any changes can be done at any step, in a transparent and explicit way, therefore agreeing with the open science principles.

Conclusions

This work presents an open-source geospatial framework that allows straightforward assessment and analysis of beach litter. The framework is based on FOSS, with GIS and statistical interfaces, being flexible and easily adapted to any beach and litter-related scenario. The procedure brings to users comprehensive baseline graphical and descriptive products that can support managers and researchers towards publication of their data, with up-to-date analytical products. This framework enables effective monitoring, produces fast results, and helps to identify beach litter accumulations and temporal trends. The information obtained can also support management strategies orientated to reduce or at least minimize the polluting activities that contribute most to litter generation along beaches. The use of FOSS facilitates the adaptation of techniques to specific realities, dispensing the need for the large financial investments associated with proprietary geospatial software. This can increase the accessibility of monitoring instruments for developing regions, to better understand the litter issue worldwide. Advantages of the proposed framework are that it is based on newly updated monitoring data and methodologies, as well as its flexibility and adaptability to the specific features of each site to be investigated, without losing data standards and being transparent in the entire process.

Acknowledgments Thanks to Andrea Sanchez for her kind review of the R code and Marco C. Brustolin for reviewing an early version of this manuscript. NWD received a Master's Scholarship from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), through the Programa de Pós-Graduação em Oceanografia Biológica (FURG). MPSM received a Master's Scholarship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), through the Programa de Pós-Graduação em Oceanografia (UFSC). JB's research activities were partially sponsored by CNPq (grant 307797/2016-3). This joint contribution was developed between Laboratório de Oceanografia Costeira (UFSC, Brazil) and the Geology, Geophysics, and Marine-Coastal Process Group (Universidad del Atlántico, Colombia). We would like to thank two anonymous reviewers who greatly improved the quality of the manuscript.

Availability of data and material https://raw.githubusercontent.com/JesSchattschneider/litter/master/litter_all.csv https://raw.githubusercontent.com/JesSchattschneider/litter/master/survey03.csv

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Code availability The R code is available in Online Resource 3.

References

Araújo, M. C. B., Santos, P. J., & Costa, M. F. (2006). Ideal width of transects for monitoring source-related categories of plastics on beaches. *Marine Pollution Bulletin, 52*, 957–961. https://doi.org/10.1016/j.marpolbul.2006.04.008.

Araújo, M. C. B., Silva-Cavalcanti, J. S., & Costa, M. F. (2018). Anthropogenic litter on beaches with different levels of development and use: a snapshot of a coast in Pernambuco (Brazil). *Frontiers in Marine Science, 5*, 233. https://doi.org/10.3389/fmars.2018.00233.

Baker, M. (2016). Is there a reproducibility crisis? A Nature survey lifts the lid on how researchers view the crisis rocking science and what they think will help. *Nature, 533*, 452–455.

Bennett-Martin, P., Visaggi, C. C., & Hawthorne, T. L. (2016). Mapping marine debris across coastal communities in Belize: developing a baseline for understanding the distribution of litter on beaches using geographic information systems. *Environmental Monitoring and Assessment, 188*, 557. https://doi.org/10.1007/s10661-016-5544-4.

Broman, K. W., & Woo, K. H. (2018). Data organization in spreadsheets. *The American Statistician, 72*, 2–10. https://doi.org/10.1080/00031305.2017.1375989.

Collins, S., Genova, F., Harrower, N., Hodson, S., Jones, S., Laaksonen, L., Mietchen, D., Petrauskaitė, R., & Wittenburg, P. (2018). *Turning FAIR data into reality – final report and action plan from the European Commission Expert Group on FAIR Data*. Luxembourg: Publications Office of the European Union https://hdl.handle.net/20.500.12259/103794.

Corraini, N. R., Lima, A. S., Bonetti, J., & Rangel-Buitrago, N. (2018). Litter and its scenic impact on the North Santa Catarina island beaches, Brazil. *Marine Pollution Bulletin, 131*, 572–579. https://doi.org/10.1016/j.marpolbul.2018.04.061.

Cowger, W., Gray, A. B., & Schultz, R. C. (2019). Anthropogenic litter cleanups in Iowa riparian areas reveal the importance of near-stream and watershed scale land use. *Environmental Pollution, 250*, 981–989. https://doi.org/10.1016/j.envpol.2019.04.052.

Derraik, J. G. B. (2002). The pollution of the marine environment by plastic debris: a review. *Marine Pollution Bulletin, 44*, 842–852. https://doi.org/10.1016/S0025-326X(02)00220-5.

Dickinson, J. L., Shirk, J., Bonter, R., Crain, R. L., Martin, J., Phillips, T., & Purcell, K. (2012). The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment, 10*, 291–297. https://doi.org/10.1890/110236.

Galgani, F., Hanke, G., & Maes, T. (2015). Global distribution, composition, and abundance of marine litter. In M. Bergmann, L. Gutow, & M. Klages (Eds.), *Marine Anthropogenic Litter* (pp. 29–56). New York: Springer.

Garcia-Silva, A., Gomez-Perez, J. M., Palma, R., Krystek, M., Mantovani, S., Foglini, F., Grande, V., De Leo, F., Salvi, S., Trasatti, E., Romaniello, V., Albani, M., Silvagni, C., Leone, R., Marelli, F., Albani, S., Lazzarini, M., Napier, H. J., Glaves, H. M., Aldridge, T., Meertens, C., Boler, F., Loescher, H. W., Laney, C., Genazzio, M. A., Crawl, D., & Altintas, I. (2019). Enabling FAIR research in Earth Science through research objects. *Future Generation Computer Systems, 98*, 550–564. https://doi.org/10.1016/j.future.2019.03.046.

Gouveia, C., Foseca, A., Câmara, A., & Ferreira, F. (2004). Promoting the use of environmental data collected by concerned citizens through information and communication technologies. *Journal of Environmental Management, 71*, 135–154. https://doi.org/10.1016/j.jenvman.2004.01.009.

Haarr, M. L., Westerveld, L., Fabres, J., Iversen, K. R., & Busch, K. E. T. (2019). A novel GIS-based tool for predicting coastal litter accumulation and optimising coastal cleanup actions. *Marine Pollution Bulletin, 139*, 117–126. https://doi.org/10.1016/j.marpolbul.2018.12.025.

Hastings, E., & Potts, T. (2013). Marine litter: Progress in developing an integrated policy approach in Scotland. *Marine Policy, 42*, 49–55. https://doi.org/10.1016/j.marpol.2013.01.024.

Hyder, K., Townhill, B., Anderson, L. G., Delany, J., & Pinnegar, J. K. (2015). Can citizen science contribute to the evidence-base that underpins marine policy? *Marine Policy, 59*, 112–120. https://doi.org/10.1016/j.marpol.2015.04.022.

Jambeck, J. R., & Johnsen, K. (2015). Citizen-based litter and marine debris data collection and mapping. *Computing in Science and Engineering, 17*, 20–26. https://doi.org/10.1109/MCSE.2015.67.

Martin, C., Parkes, S., Zhang, Q., Zhang, X., McCabe, M. F., & Duarte, C. M. (2018). Use of unmanned aerial vehicles for efficient beach litter monitoring. *Marine Pollution Bulletin, 131*, 662–673. https://doi.org/10.1016/j.marpolbul.2018.04.045.

Moreira, F. T., Balthazar-Silva, D., Barbosa, L., & Turra, A. (2016). Revealing accumulation zones of plastic pellets in sandy beaches. *Environmental Pollution, 218*, 313–321. https://doi.org/10.1016/j.envpol.2016.07.006.

Newman, G., Wiggins, A., Crall, A., Graham, E., Newman, S., & Crowston, K. (2012). The future of citizen-science: emerging technologies and shifting paradigms. *Frontiers in Ecology and the Environment, 10*, 298–304. https://doi.org/10.1890/110294.

Nunes, M., Ferreira, Ó., Schaefer, M., Clifton, J., Baily, B., Moura, D., & Loureiro, C. (2009). Hazard assessment in rock cliffs at Central Algarve (Portugal): a tool for coastal management. *Ocean & Coastal Management, 52*, 506–515. https://doi.org/10.1016/j.ocecoaman.2009.08.004.

Ocean Conservancy. (2010). *Trash travels – 2010 annual report*. Washington, DC: Ocean Conservancy https://oceanconservancy.org/wp-content/uploads/2017/04/2010-Ocean-Conservancy-ICC-Report.pdf. Assessed 15 December 2018.

QGIS Development Team. (2018). *QGIS Geographic Information System*. Chicago: Open Source Geospatial Foundation https://www.qgis.org.

R Core Team. (2019). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing https://www.R-project.org.

Rangel-Buitrago, N., Williams, A., Anfuso, G., Arias, M., & Gracia, C. A. (2017). Magnitudes, sources, and management of beach litter along the Atlantico department coastline, Caribbean coast of Colombia. *Ocean & Coastal Management, 138*, 142–157. https://doi.org/10.1016/j.ocecoaman.2017.01.021.

Ryan, P. G., Moore, C. J., van Franeker, J. A., & Moloney, C. L. (2009). Monitoring the abundance of plastic debris in the marine environment. *Philosophical Transactions of the Royal Society B, 364*, 1999–2012. https://doi.org/10.1098/rstb.2008.0207.

Santos, I. R., Friedrich, A. C., Wallner-Kersanach, M., & Fillmann, G. (2005). Influence of socio-economic characteristics of beach users on litter generation. *Ocean & Coastal Management, 48*, 742–752. https://doi.org/10.1016/j.ocecoaman.2005.08.006.

Schneider, F., Parsons, S., Clift, S., Stolte, A., & McManus, M. C. (2018). Collected marine litter—a growing waste challenge. *Marine Pollution Bulletin, 128*, 162–174. https://doi.org/10.1016/j.marpolbul.2018.01.011.

Schulz, M., Walvoort, D. J. J., Barry, J., Fleet, D. M., & van Loon, W. M. G. M. (2019). Baseline and power analysis for assessment of beach litter reductions in the European OSPAR region. *Environmental Pollution, 248*, 555–564. https://doi.org/10.1016/j.envpol.2019.02.030.

Seeliger, U., Odebrecht, C., & Castello, J. P. (1997). *Subtropical convergence environments: the coast and the sea in the Southwestern Atlantic*. Berlin: Springer.

Tanhua, T., Pouliquen, S., Hausman, J., O'Brien, K., Bricher, P., Bruin, T., Buck, J. J. H., Burger, E. F., Carval, T., Casey, K. S., Diggs, S., Giorgetti, A., Glaves, H., Harscoat, V., Kinkade, D., Muelbert, J. H., Novellino, A., Pfeil, B., Pulsifer, P. L., Van de Putte, A., Robinson, E., Schaap, D., Smirnov, A., Smith, N., Snowden, D., Spears, T., Stall, S., Tacoma, M., Thijsse, P., Tronstad, S., Vandenberghe, T., Wengren, M., Wyborn, L., & Zhao, Z. (2019). Ocean FAIR data services. *Frontiers in Marine Science, 6*, 440. https://doi.org/10.3389/fmars.2019.00440.

Topçu, E. N., Tonay, A. M., Dede, A., Öztürk, A. A., & Öztürk, B. (2013). Origin and abundance of marine litter along sandy beaches of the Turkish Western Black Sea Coast. *Marine Environmental Research, 85*, 21–28. https://doi.org/10.1016/j.marenvres.2012.12.006.

Velander, K., & Mocogni, M. (1999). Beach litter sampling strategies: is there a "best" method? *Marine Pollution Bulletin, 38*, 1134–1140. https://doi.org/10.1016/S0025-326X(99)00143-5.

Whickham, H., & Grolemund, G. (2017). *R for data science*. Sebastopol: O'Reilly Media.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Bomberg, N., Boiten, J.-W., Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data, 3*, 160018. https://doi.org/10.1038/sdata.2016.18.

Williams, A. T., & Rangel-Buitrago, N. (2019). Marine litter: solutions for a major environmental problem. *Journal of Coastal Research, 35*, 648–663. https://doi.org/10.2112/JCOASTRES-D-18-00096.1.

Zettler, E. R., Takada, H., Monteleone, B., Mallos, N., Eriksen, M., & Amaral-Zettler, L. A. (2017). Incorporation citizen science to study plastics in the environment. *Analytical Methods, 9*, 1392–1403. https://doi.org/10.1039/C6AY02716D.