ROYAL SOCIETY
OF CHEMISTRY

PAPER

Check for updates

Intersections between materials science and marine plastics to address environmental degradation drivers: a machine learning approach†

Henrique de Medeiros Back, [ID] *[a] Daphiny Pottmaier,[b] Camilla Kneubl Andreusi [ID] [a] and Orestes Estevam Alarcon[a]

Plastics are an integral part of the material structure in modern societies. However, their widespread contamination in the environment raises concerns regarding the sustainable use of these materials. Plastic pollution research has accelerated rapidly in the past 20 years and developed as a broad and multidisciplinary field, due to its complex nature. Research on the fate of plastics in the environment, specifically their dispersal and degradation, is one of the pillars in the field. In this study, we used machine learning and text mining tools to bridge the marine plastics community to the existing knowledge from polymer science regarding degradation drivers. Topic modelling enabled visualising hot topic trends in marine plastics research and filtering for relevant publications with minimal expert intervention. The recurrence of drivers was verified in the literature, indicating particular areas of focus on the marine degradation of plastics. The results show weathering conditions to be more recurrent than polymer properties and the latter to be rarely discussed in depth. Also, biodegradation is found to be a hot topic in the field, while degradation caused by abiotic factors is less addressed. This may be necessary as polymer engineering is traditionally less concerned about the end-of-life of plastic products. Finally, we argue that not only a deep understanding of plastics from polymer science should aid fundamental degradation studies by the marine plastics community, but also that the latter has the opportunity to largely contribute to the former by filling the gaps it has left.

Environmental significance

Plastics are ubiquitous in everyday life and the environment. Recently, the academic community has put tremendous effort into trying to describe its causes and understand its consequences. Notably, the fates of plastic particles are of particular interest. This has led to a rapidly growing repository of new publications that even experts struggle to be up to date. Machine learning can be a powerful tool to help synthesize and filter through these large amounts of unstructured textual data dispersed in scientific publications. This paper uses text mining to semi-automatically filter the relevant literature, explore trends in the field and verify possible mismatches between polymer science and the environmental plastics community. Our results show essential features of degradation not receiving due attention in the more recent marine plastics research. Topic trends also reveal hot topics in the field and the impact of the COVID-19 pandemic. These allow a broad view of how the field is evolving and of the relative importance of particular topics while showing the usefulness of machine learning for the apprehension of rapidly growing scientific pieces of literature.

Introduction

Plastic production has grown exponentially since the end of World War II and they are now used in all sectors of industry and day-to-day applications, replacing a multitude of other materials, especially natural ones. Since then, they have become so important to modern life that the current period of human history has been described as the Plastic Age, in relation to the Bronze and Iron Ages.[1–3] The accumulation of plastic waste in nature, especially in the oceans, has been a topic of great public and scientific interest in recent years. Even if a more thorough understanding of impacts that range from entanglement, ingestion, and toxicity is still developing, the scale of this issue demands precautionary action.

Plastic pollution research typically focuses on three areas: sources, fates, and impacts.[4] The fates of plastic particles in different environments are influenced by their behaviours, such as horizontal and vertical transport, weathering and fragmentation, particle aggregation, biofouling, and ingestion.[5,6] New findings in these areas help identify accumulation hotspots, removal processes, regional/global mass budgets, degradation/fragmentation rates, and affected areas/species. Knowledge of

[a]*Universidade Federal de Santa Catarina, Brazil. E-mail: henrique.m.back@posgrad.ufsc.br*

[b]*Nottingham Trent University, UK*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d3va00106g

© 2023 The Author(s). Published by the Royal Society of Chemistry

Environ. Sci.: Adv., 2023, **2**, 1629–1640 | **1629**

particle concentrations and characteristics is crucial in understanding how plastic pollution affects ecosystems. The issue of plastic pollution is also being studied from the perspective of social and political sciences, ocean physics, and chemistry. Particularly regarding the behaviour and fate of plastics in the environment, chemical and materials engineering can contribute with knowledge of polymer structures, their properties and mechanisms that lead to their degradation and change in characteristics.

Reviews provide mechanisms to synthesise the literature and generate new insights, but traditional systematic reviews are time-consuming and subjective. In a fast-growing scientific field, by the time the review is published, it is already outdated. A new era in research calls for new methods of analysis and artificial intelligence (AI) may be our way to adapt. The use of AI and machine learning (ML) has grown in plastic pollution research and other environmental fields as tools to analyze large amounts of data, create predictive models, and reduce human error. Natural language processing (NLP) allows the transformation of unstructured text into data that can feed ML models. These versatile tools can help researchers navigate and analyze the growing flood of information.

This paper aims to visualize trends in plastic pollution research and, specifically, how often particular terms associated with the degradation of plastics are being considered. We used text mining and BERTopic, a topic modeling software, for this purpose. We show the focus given by the community on the plastics fate problem and point out the aspects of degradation that are being left out and should be considered.

Methods

A workflow of the methods is shown in Fig. 1. To assess how polymer properties are related to materials degradation in the environment, we draw from the polymer science literature and briefly explain some concepts in the Degradation impact drivers section (more detailed explanations can be found in Appendix A). In the Results section we present a table summarizing the linkages between drivers and degradation processes as having high, moderate or low relevance. This is done similarly to how the IPCC's Sixth Assessment Report presents the relevance of key climatic impact drivers. 'High relevance' indicates drivers that are most prominent for their direct connection to degradation processes, while 'lower relevance' indicates weaker linkages or driving behaviours.

To understand how this knowledge is being incorporated by the marine plastics community, we counted the recurrence of key terms in the abstracts of this literature. The construction of a vocabulary for text mining was made concurrently with the identification of drivers, to find compatible terms between both literature studies. Only abstracts were used because they are readily available as metadata on SCOPUS and are summaries of the full text. They are more commonly used, but full texts may be employed, especially if there is a smaller literature available.

Papers were sourced from Scopus using the following search query, adapted from Ivar do Sul, 2018:[7] "TITLE-ABS-KEY (("marine debris" OR "marine plastic debris" OR "marine plastic pollution" OR "plastic pollution" OR "marine plastic" OR "plastic debris" OR microplastic OR nanoplastic OR microfiber OR "plastic pellet") AND (marine OR coastal OR sea OR ocean OR beach)). The search was performed on February 24, 2022, and resulted in 8319 publications. Then, all abstracts were retrieved alongside the corresponding year of publication and DOI. We included a step to remove duplicates based on DOI identifiers as well as instances without abstracts.

For topic modelling, documents were embedded using SPECTER, a pre-trained BERT language model that is trained on scientific text.[8] This was preferred over the standard MiniLM used by BERTopic, as it should be more capable of capturing semantic relationships between words in this kind of text. Indeed, initial trials with both models showed more favourable results with SPECTER. Dimensionality reduction and clustering steps were defined using the standard BERTopic framework. It consists of performing these operations using UMAP (uniform manifold approximation and projection) and HDBSCAN (hierarchical density-based spatial clustering of applications with noise), respectively. The seed of randomness was fixed at "42" (this choice is arbitrary) to allow the reproducibility of results since the modelling is a stochastic process. It's important to note that the method utilises an unsupervised learning approach, which means we have no *a priori* knowledge of the topics present in our dataset. For further details refer to Appendix A.

A topic model was generated using BERTopic over the embedded text. It attributes each document to a topic based on the calculated probabilities of each document belonging to each topic. However, if none of these is higher than a minimum threshold, the model assigns them to an outlier topic. An initial analysis of this topic's representation indicated that many documents relevant to our study were set to be outliers. To avoid that, the threshold was set to 0.01 – after also having checked the topic probability distribution for several documents and tried using 0.02 and 0.005, which, respectively, left thousands and dozens of outliers. This was made with the assumption that the literature was well represented by all generated topics and that most documents would belong to at least one of them.

To evaluate the topic modelling, we chose the silhouette score as it is intuitive and simple to implement. It is a measure of the similarity between a document and its own topic in relation to its difference from other topics. Scores can range from -1 to 1, where 1 means topics are cohesive and well separated. However, it should be kept in mind when evaluating the model that all evaluation metrics are only indicators of a model's performance and have better use in benchmarking different algorithms. Their evaluation and utility are ultimately dependent on the user's subjectivity.

To be able to identify what topics are about, we created topic representations using vectorization and c-TF-IDF (see Appendix A for further information). During vectorization, the stop_word parameter was set to "English" to exclude words that add no meaning to the text (*i.e.*, and, the, in, *etc.*) and could hinder topic differentiation. Another parameter, "ngram_range" was set as (1, 2) to include compound words like "plastic waste" instead of looking just at "plastic" and "waste".

Fig. 1 Workflow of the methods for vocabulary construction and text mining.

By analysing the hierarchy of topics, we were able to remove topics unrelated to the scope of the study and merge topics that were closely related. This step required a more in-depth look into topics by inspecting topic word scores, that is, how much each word is associated with a certain topic, and representative docs, which are the abstracts that better represent a topic (all of this can be seen in detail in the intersections.ipynb file in the ESI†). This step (merging and removing) was conducted recursively two other times as topic representations are updated and may better reveal the meaning of certain topics.

To visualize the development of the field, the number of publications was plotted per year and by topic. We used the "topics over time" function on BERTopic to calculate the frequency of papers present in each topic over the studied period. Topic representations are dynamic which allows us to see how topics evolve as well (interactive plot available in the intersections.ipynb file in the ESI†).

After the modelling and screening of papers, keyword recurrence was verified in the corpus of each topic. For this, we defined a function in Python and applied it to our corpus using the constructed vocabulary. The output recurrence is the number of abstracts that contain at least one mention of a particular word, which is not the same as word occurrence which would count every mention of that word. Details can be found in the text_mining.ipynb file in the ESI.†

Degradation impact drivers

Degradation is any form of loss of function or property impaired by some kind of agent as a result of physicochemical modifications in the material. In other terms, degradation imparts changes in composition and/or structure that affect properties. This relationship between a material's composition, microstructure, processing, and properties is widely accepted in materials science. A conventional way to represent this is shown in Fig. 2a. On a molecular level,‡ any degradation of polymers starts with the breaking of a covalent bond in the macromolecule, which generates reactive species, called free radicals, responsible for propagating the degradation reaction. This energy can be provided through heat, light, chemical potential or mechanical stress (Fig. 2b). The magnitude of the energy needed to break the bond depends on the material properties. Therefore, we define degradation drivers as material properties and environmental factors that interact to cause the weathering of plastics.

Material characteristics

Composition

The terms "plastic" and "polymer" are often used interchangeably in the environmental literature, which is not accurate. Plastics are materials that belong to the class of polymers, which has a unifying factor in their chemical structure: all polymers form long chains from the repetition of a monomeric unit. They can occur naturally, like proteins or cellulose, or be synthesized artificially. The latter are commonly referred to as

---

‡ Plastic degradation can occur without molecular changes, as in the case of additive desorption or water adsorption and contribute to loss of mechanical resistance, thus increasing the fragmentation rate. However, these were not discussed.
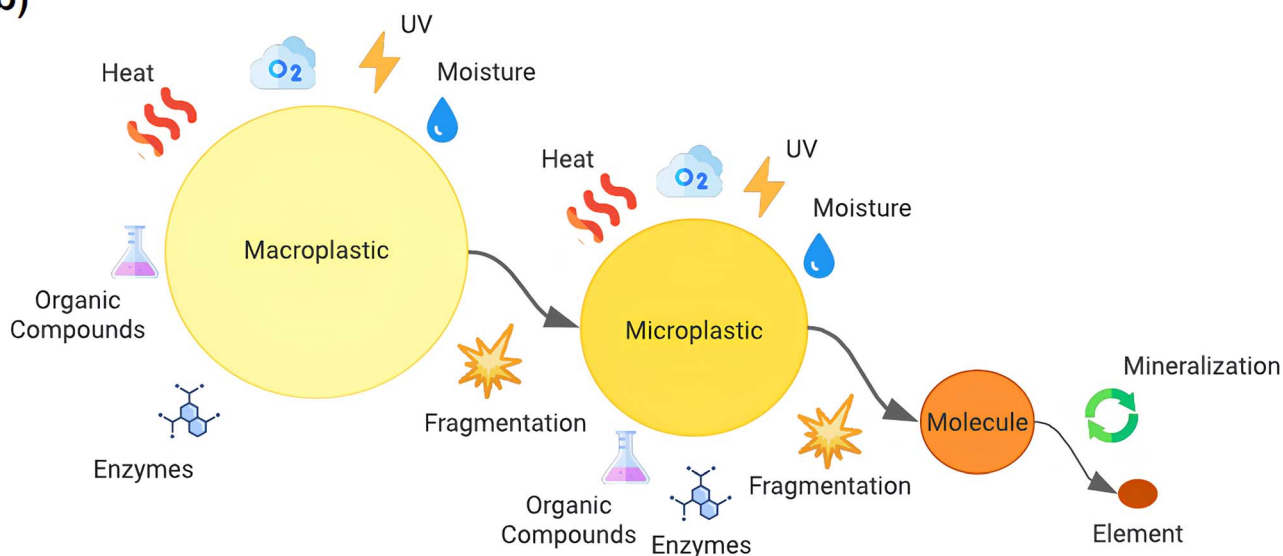
Fig. 2 (a) The materials science tetrahedron; (b) particle size reduction as caused by degradation and fragmentation of plastics.

"plastics", due to the plasticity that many of them have – especially under the application of heat, a characteristic that differentiates them from the majority of natural polymers. Considering the varied mixture of substances (additives) in plastics, the term "polymer" can also be used when referring to the main resin that makes up a plastic product/particle or simply to its macromolecular chains.

These macromolecules consist mostly of C–C (carbon–carbon), C–H (carbon–hydrogen) and C–R (carbon–radical) bonds where R can be a halogen, a functional group, such as an aromatic ring, *etc.*[10,11] The energy of bonds depends on their nature and that of the surrounding bonds in the molecule. C–C bonds can also be of different types, or hybridised states, as shown in Fig. 3. Primary carbons will have higher C–H bond energy and will be less prone

to bond dissociation. Tertiary carbons are more susceptible to bond breaking in the C–H bond than the others because that carbon is already strongly bonded to other carbons.[12] The same effect is observed for C–H bonds surrounded by other chemical species, such as C=C (polystyrene), C–O–C (like in polyesters), and C–N–C (like in polyamides) that have stronger bond energies.[13] Fig. 3 illustrates a section of a polymer chain with different carbons and other structures mentioned further in the text and provides typical bond energy for several species. The actual values can only be stated in relation to the surrounding chemical structure. Softwares like Gaussian and Gamess can be used to predict actual bond energies in complex molecules.[15] Generally, strong bonds will reduce the energy of surrounding bonds.
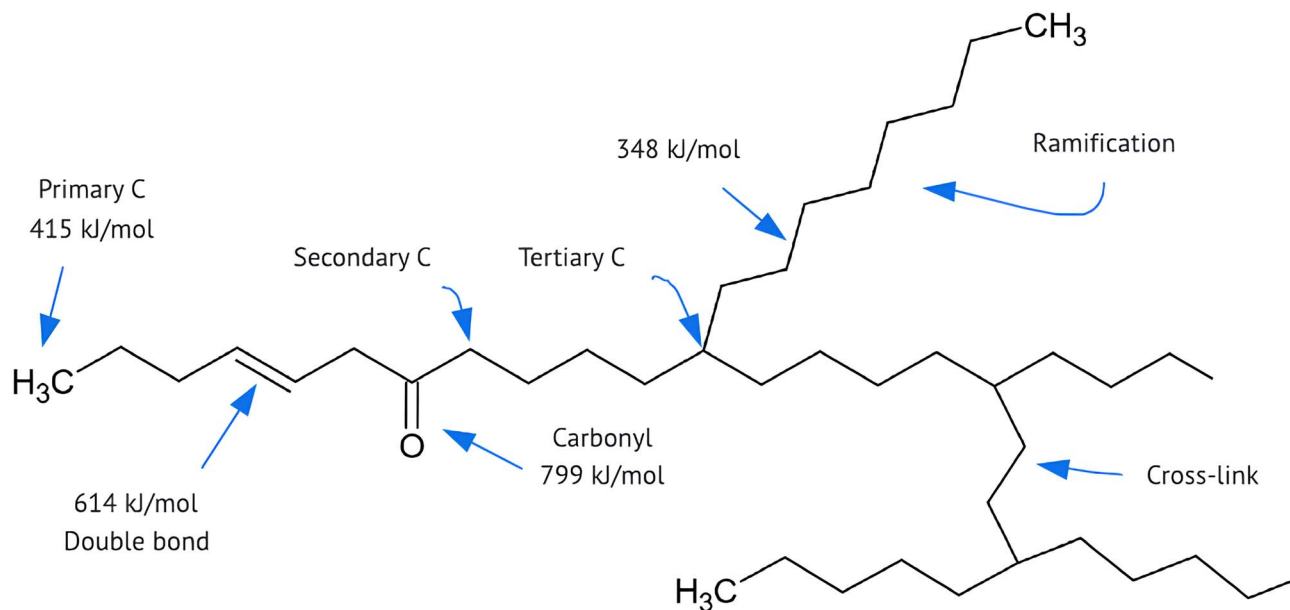
Fig. 3 Example section of a polyethylene macromolecule with possible deviations from the ideal structure and corresponding dissociation energy of some typical bonds. Values are related to the bonds made between carbon and carbon or hydrogen or oxygen. Source of values: Huheey, Keiter and Keiter, 1993.[14]

Microstructure

When studying polymer degradation, another point that can be broadly addressed is the strength of intermolecular forces, as its increase will impart greater stability towards degradation.[16] Stronger intermolecular forces can be imparted by strong polar groups or by closely packed macromolecules. The first is directly related to chemical composition (already addressed), while the second is related to the physical microstructure of the polymeric material.

Polymers' macromolecules have a high molecular weight (MW), and for that, they will behave very differently compared to the substances containing just the monomers. Increasing MW increases the intermolecular forces between the polymer chains, which is related to the cohesive energy responsible for the mechanical strength and chemical stability of polymers.[9] MW is always in some range and this distribution is also a factor of importance for properties.

Polymers can also be amorphous or semi-crystalline. The degree of crystallinity, or the ratio between crystalline and amorphous phases in the material, affects the strength of intermolecular forces because crystals have a more compact microstructure. The closer these molecules are, the more they will interact with each other. This increases mechanical strength, melting temperature and stability towards oxidation and many other degrading agents.[17,18] This includes stability to biodegradation, as we will see further in the text. Engineers control this ratio to produce materials with the desired properties, by controlling processing conditions, MW and composition, especially bond rigidity and side group sizes.

A similar effect is the result of cross-linking, which is the formation of covalent bonds between polymer chains that can happen during polymerization or as a result of degradation.[19] It

reduces molecular mobility, because it ties one molecule to the other, restricting their conformation, thus increasing bulk stiffness and elevating the glass transition temperature (T_g) and melting temperature (T_m). Contrarily, branching, or ramifications, which are secondary polymer chains attached to a primary backbone, weaken intermolecular forces because they reduce densification and impair crystallization. Additionally, they weaken intramolecular forces through the presence of tertiary carbons.[20]

Some other considerations can be made regarding physical arrangements and the microstructure, such as the form of crystals or the degree and length of ramifications, but they will not be addressed here for brevity. In any case, they will also influence chain rigidity or crystallinity.[21]

Properties

Plastic properties are the macroscopic characteristics of the material that determine its behaviour during use or after end-of-life. Degradability, or even compostability, for instance, are some of these properties. As stated before and shown in Fig. 2a, properties are conditioned to the composition and microstructure. However, they are more convenient to measure. A good example is the glass transition temperature, a typical property of amorphous or semicrystalline materials that is commonly measured for polymers. It indicates the mobility capacity of polymeric chains and it's correlated to several other properties, such as molar mass, degree of crosslinking and crystallinity.[21,22] While it is not involved in the degradation process, it may be a relevant measure as it reflects composition and structure factors that are.

Likewise, density can also be of interest. Determining whether a particle will float or sink away from the surface of

a waterbody and away from sun exposure influences the intensity of degradation. Mechanical strength, on the other hand, is important for fragmentation and when reduced will tend to increase the surface-area-to-volume ratio, but it is not directly related to degradation itself. However, it is a useful measure of the degree to which a plastic particle has been degraded already and can thus, appear in the literature as such.

Weathering factors

Ultraviolet radiation

Photons of ultraviolet radiation can initiate a photochemical reaction that causes degradation. Once started, it will propagate due to the presence of free radicals. It requires a chemical species to be in an excited electronic state, which happens when an unbound electron increases its energy by absorbing a photon, mostly in the ultraviolet (295–380 nm) range.[23] The mere incidence of light, however, does not necessarily cause excitation. It must be absorbed. And since energy is quantized, a chemical bond will only absorb photons at specific frequencies. The chemical group responsible for light absorption in the UV-Vis region is called a chromophore. It can give colour to a material since only part of the incident radiation will be reflected but is also responsible for photodegradation. There are two types of chromophores: intrinsic (present in the macromolecule) and extrinsic (additives or contaminants). Photostabilizers are additives used in plastics production as protection from photodegradation because they contain chromophore groups that preferentially absorb UV.

Most polymers have no intrinsic chromophore groups but can degrade under UV because of impurities or additives. PE, for instance, can undergo photochemical degradation, because in real-world conditions there will necessarily be contaminants or previous degradation that allow it. Similarly, PS does not absorb at any frequencies in the solar spectrum, but irradiation can generate an excimer (association of two adjacent aromatic rings). This excited state is responsible for initiating the photodegradation process even in energy bands not absorbed by chromophores. In most polymers, these defects will be present in the form of C=C conjugated or aromatic double bonds or even as carbonyl groups (C=O).

Photodegradation will cause the appearance of carbonyl groups (rigid double bonds) and crosslinking, both reducing the mobility of polymeric chains in the plastic material. However, since the yield of chain scission depends on chain mobility, the process will slow down with further degradation[24] Additionally, photodegradation only occurs on the surface of the material up to a few micrometres, depending on the energy of the photon and the material.[25]

Temperature

Due to their relatively weak covalent bonds, the maximum service temperature of polymers is considered low if compared to other materials, like most metals and ceramics, ranging typically from 50 °C to 200 °C.[28] The effect of temperature on degradation cannot be neglected though, as it influences many other chemical processes, such as migration of low molecular weight additives, hydrolysis and photo-oxidation.[26,27,29] Due also to their thermal insulation capacity, polymers will suffer thermal degradation differently depending on how long they are exposed to higher temperatures.[30]

Polymers with the same types of covalent bonds will have different thermal decomposition temperatures due to molecular arrangements. For instance, PE and PP only have C–C and C–H bonds, however, PE will only contain low-energy tertiary carbons in its ramifications, while PP has one of such carbons at each monomeric repetition. This makes PP less stable and more prone to oxidation.[26]

Once thermal degradation is initiated it is highly favoured by free radicals and other compounds typically found in polymers, such as residual monomers and catalysts, as well as defects. These may originate from the production process itself, either during polymerization or from heating or chemical contamination during manufacturing. Additionally, chain defects caused by previous degradation can also reduce thermal stability.[31] In contrast to photodegradation, thermal degradation occurs in the bulk of the material, not just at the surface.

Chemicals

Polymer degradation can also occur by reactions with chemicals. It can happen at the surface of the material due to contact with substances in the environment, or in the bulk, caused by impurities or by the diffusion of liquids, such as water. Unlike UV or temperature, chemicals can cause depolymerization, where the polymer is converted back into monomers, which is good for recycling but is a source of toxic compounds in the environment.

The reactivity depends on the affinity of both polymer and chemical. Therefore, it should be analysed on a case-by-case basis. However, shared characteristics of substances can be indicators of reactivity. For instance, it will be higher for organic compounds than for inorganic salts, alkaline or weak acid solutions. Additionally, more hydrophobic materials, such as polyolefins, will tend to react more intensely with oils and other hydrophobic compounds. Persistent organic pollutants are some of these compounds that co-occur with plastics in marine environments and can favour their degradation. Similarly, polymers containing polar groups such as polyamides, polyesters and polycarbonates, have higher interactions with water that can cause hydrolysis.

Microbes

The general process of biodegradation involves the excretion of enzymes by microorganisms that cause hydrolysis of the polymer chain into intermediate compounds that can then be assimilated by these organisms as a source of carbon.[32] Several polymers have been reported to suffer biodegradation, mostly from bacteria or fungi,[33] but polymers containing ester, amide or carbonate bonds in the backbone are more prone to it, because of their susceptibility to hydrolysis in these bonds.

To produce biodegradation, microbes must have access to water and the polymer chain. This is generally related to how

compact a certain microstructure is and how easy it is to rotate macromolecules around molecular bonds. Strong bonds, such as C=C and C=O, have the tendency to reduce the energy of adjacent bonds, which makes rotating around them easier.[34] These carbonyl bonds also increase hydrophilicity, which facilitates microbe adhesion and water availability. Additionally, bulky side groups, ramifications and cross-linking all limit the molecules' conformability, thus hindering biodegradation.[34] The degree of crystallinity also has a role, as the crystalline structure is more compact and less free to move. The molecular weight is of particular importance, as microbes cannot assimilate long chains. Even polymers considered biodegradable, such as PLA and PHB may have to suffer previous chain scission in order to be metabolized by microbes.[35]

Aside from polymer properties, environmental conditions also have an influence on biodegradation. Temperature is the first factor to be considered. Increases in temperature increase the rate of biodegradation, until an optimum value above which biological activity is reduced. Generally, thermophilic composting (up to 70 °C) in institutional facilities produces faster biodegradation than mesophilic home compost (around 35 °C).[36] Moisture, pH, oxygen and UV radiation are also decisive factors for microbial growth and will influence the rate of biodegradation.[37]

Results & discussion

Fig. 4 shows the linkages between key degradation drivers and degradation processes and compartments summarized from the previous sections where we can easily identify drivers with high relevance for one or more degradation processes. These linkages will be considered when we evaluate the recurrence of drivers in the marine plastics literature. But first, we classify the publications in this literature by performing our topic modelling with BERTopic as described in the Methods section. A total of 7828 abstracts from the preprocessed dataset based on our search query on SCOPUS were included. Initially, the modelling resulted in 89 topics and around 3000 documents were assigned to an outlier topic. After applying a new threshold, only 41 documents were classified as outliers.

It is important to note that the topics found can differ using other hyperparameters, models or even slightly different datasets.[38] In any case, the modelling gives us enough information about topics contained in our dataset, that we can use to understand it. Despite being able to distinguish between very specific topics, the great number of topics given by the model hindered a better comprehension of what is going on in the literature. So, topics were merged by similarity as described in the Methods section. Additionally, we removed a total of 6 topics consisting of 242 abstracts (3%) that were outside the scope of our analysis in topics such as "facies_sandstone_deposits_mass transport_mounds", "sensor_sensitivity_salinity_microfiber_high   sensitivity", "composites_nanofibers_microfibers_melt_collagen"   and "soil_soils_microplastics soil_agricultural_terrestrial". Finally, the literature could be represented by 41 different topics.

The calculation of the silhouette score for the evaluation of topic cohesion resulted in a value of 0.26. With values ranging from −1 to 1, a score of 0.26 is reasonable. While we cannot be certain if the topic assignment of documents is correct and we could be leaving out papers that should be considered, the silhouette score, accompanied by subsequent topic explorations, is sufficient for enough confidence in the results. Moreover, the method drastically reduces individual subjectivity which is concerning when screening the literature.

In Fig. 5 we show the frequency of new publications per topic per year since 2010. The period was only chosen to facilitate the visualisation, as frequencies remained relatively constant before that. Similarly, not every topic was selected to facilitate visualization. The internal build of the algorithm allows the user to visualise not only the frequency of documents per topic per year but also how each topic representation changes over time. We have found that they don't change significantly, but sometimes a word may appear in a topic's representation, indicating that it has gained relevance in that topic.

From Fig. 5, we observe that "topic 1 – ingestion/entanglement" and "topic 4 – debris occurrence in beaches" are the most frequent until 2015, and then "topic 2 – MP in sediments" starts to grow faster and from 2017 onwards "topic 3 – exposure" has a sharp increase as well. Overall, these are the most common topics, but this result may be biased to indicate how broadly those topics were defined by the model. So we should be careful with that interpretation. However, in a general sense, the results are something that could be expected.

It may be more insightful to look at growth trends, instead. For instance, the number of publications per year in "topic 4 – occurrence in beaches" has had little change since 2020, while "topic 5 – policy" grew much faster. This may indicate a change of interest inside the research community towards new legislation instead of accounting for the occurrence of plastics. We could also say that "topic 11 – biodegradation" was a hot topic in 2022, because of the sharper increase in publications in that year. Particularly in the case of Topic 5, we can speculate the reasons for this growth as being related to key international agreements being drafted and signed, such as the plastic waste amendment to the basel convention on hazardous wastes and the UNEA resolution passed in 2022. These clearly indicate an overall interest in policy that can have an effect on scientific publications as well.

In general, all of the major topics inside the field of plastic debris in marine environments have experienced changes in their trends between 2020 and 2022. This may be explained, at least partially, by the COVID-19 pandemic. Due to lockdowns, difficulties in accessing public places and laboratories may have influenced how researchers conducted their new and ongoing projects, but also, funding and personal interests could have been shifted from some areas to others.[39,40] One example of this is the emergence, in 2020, of "topic 24 – face masks".

For some topics, you could argue that they are no longer frequent because they are saturated, with a few studies having a high citation impact and less interest in conducting new research on that topic. In Fig. 6, a look into the impact factor by topic can help validate this assumption. This may be true for

Weathering Factor	Degradation process			
	Photo	Chemical	Thermal	Bio
UV				
Temperature				
Moisture				
pH				
Oxygene				
Organic Compounds				
Microorganisms/enzymes				

Polymer properties				
Bond energy				
Linearity				
Functional groups				
Cristallinity				
Molecular weight				
Chain mobility				
Glass transition temperature				
Defects				
Chromophores		-	-	-
Carbonyl group				
Free radicals				
Residual Monomers				
Catalysts				
Additives				

Fig. 4 Linkages between degradation drivers and processes given with high, moderate and low relevance. Details on these linkages are expressed in the previous sections.

topics 7 and 9 ("drift modelling" and "persistent organic pollutants", respectively). Both have a similar slow growth curve in Fig. 4 and a higher impact factor. Additionally, documents associated with "topic 11 – biodegradation" seemed to have a smaller impact than those associated with "topic 20 – ageing" which tends to be more about physicochemical degradation.
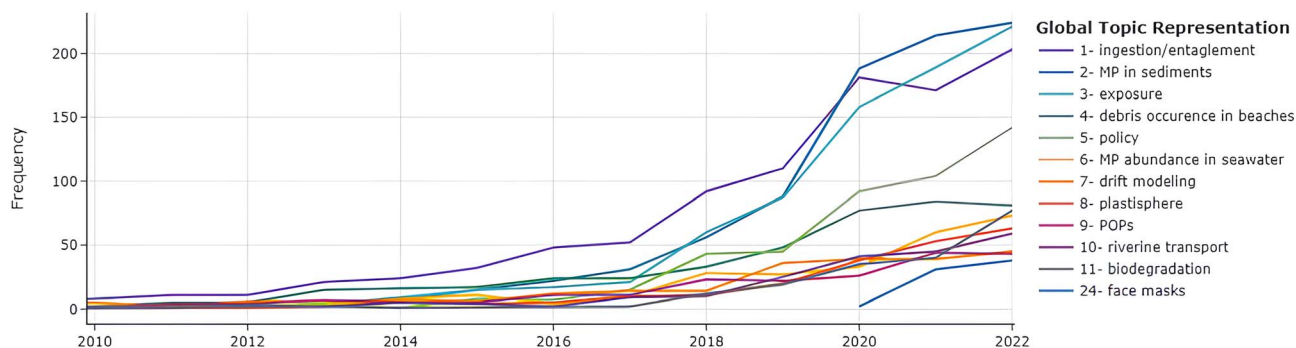
Topics over Time

Fig. 5 Topic trends in marine plastics research – some topics were omitted to facilitate visualisation.

Again, this could be because the latter is already saturated and looking at this topic's trend we see it has a small and steady growth over the years compared to a more abrupt growth in Topic 11, as mentioned previously. It could be helpful to compare "Citations over time" with "topics over time", but unfortunately, the former is not provided as metadata by SCOPUS.

Aside from topics' size and impact factor, Fig. 6 also shows the sum of normalised driver recurrence per topic. This variable accounts for the sum of recurrences of every degradation driver for each topic. This means that when the algorithm finds a word from our drivers vocabulary in an abstract it will count 1, no matter how many times that word appears in that abstract. The normalization, or division by the size of the topic, allows a better comparison between small and large topics. Interestingly, topics ""8 – plastisphere", "11 – biodegradation" and "20 – aging" have the highest recurrence of degradation drivers, which is expected and helps validate our topic modelling. Regarding Topic 8, it may not be specifically about degradation, but it certainly overlaps with it.

To gain a better understanding of which drivers are being addressed the most by each topic, we created a stacked bar graph for the normalised driver recurrence by word by topic for the top ten topics in Fig. 6, which is shown in Fig. 7.

Considering the findings shown in Fig. 4, each weathering factor can have a different impact depending on the type of degradation. By comparing them with Fig. 7, a possible interpretation is that biodegradation is being more thoroughly addressed than other modes of degradation, followed by photodegradation. This is corroborated by the fact that topics "8 – plastisphere" and "11 – biodegradation" have higher normalized sums of degradation drivers, but are also more frequent topics overall (size of bubbles in Fig. 6).

We can observe, from Fig. 7, that weathering factors are much more recurrent than material properties in topics associated with degradation except for the term "additive", which seems to be an outlier. It is relatively well mentioned in most of the selected topics, but it can be noted that its recurrence is greater in topics "15 – tire/road wear nanoplastics", "20 – ageing" and "25 – particulate/dissolved organic matter/carbon",

indicating the relevance of this property in these topics. However, due to the limitations of this analysis, it is unclear if the term "additive" is being mentioned as a potential driver of degradation or not. The same reasoning can be applied when looking at weathering factors such as oxygen, which can be present in the context of degradation or not. Mentions by topic may provide some context in the use of words. For instance, oxygen must be present in abiotic degradation to produce oxidation and formation of carbonyl bonds, so mentions in "topic 20 – ageing" are likely to be used in that context. In the context of "topic 11 – biodegradation", oxygen may be a parameter for biofouling and microorganism development. In any case, oxygen will be relevant for degradation even if indirectly.

It is pertinent to point out that "topic 20 – ageing" has a small recurrence of words like bacteria or microbe, while the recurrences for temperature, UV, and oxygen are higher. This is in good agreement with the label given to this topic which is based on the most important words for that topic (refer to Methods and Appendix for more details). Also, for material properties like crystallinity and functional groups, the recurrences are greater than in other topics. By this, we can infer that topic 20 deals more with drivers related to materials and physicochemical degradation than the other topics, especially topics 8 and 11 which are more biologically focused. The fact that our topic model was able to differentiate these topics and the finding that driver recurrences vary significantly between them show an apparent compartmentalization of studies. Although it's not possible to assert the researchers' background on any of these publications, we see that biologically focused researchers are much more likely to limit their analysis to environmental factors and dismiss material characteristics, contrarily to studies on physicochemical degradation. This result clearly points to a lack of fundamental concepts of polymer science in the environmental science community, given it is clear that both material and environment interact to produce degradation and should be equally assessed.

"Topic 38 – antibiotic resistance" contains more mentions of words like bacteria and temperature, and little to no mentions of other drivers, including enzymes. This suggests it may be
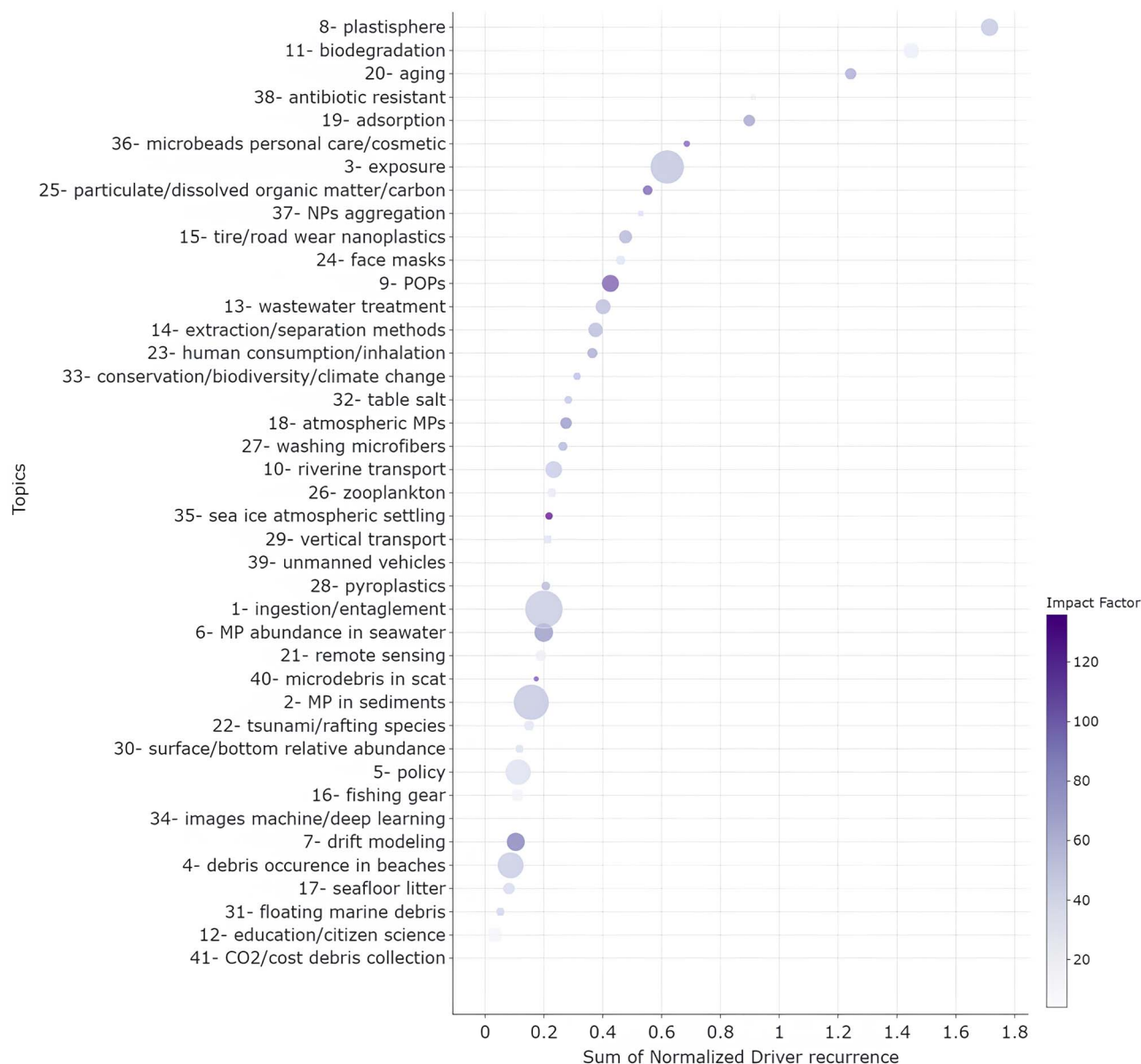
© 2023 The Author(s). Published by the Royal Society of Chemistry

Environ. Sci.: Adv., 2023, **2**, 1629–1640 | **1637**

Fig. 6 Sum of normalized recurrence of degradation drivers in the marine plastics literature by topic considering topic size (number of publications in each topic, set as size of bubbles) and topic impact factor (total number of citations for all publications in a given topic divided by the number of publications in that same topic, set as colour).

a false interpretation to consider it a relevant degradation topic and that these drivers are mentioned in other contexts. The same may be true for "topic 19 – adsorption", where "additive" and "organic pollutants" may be mentioned not as degradation drivers but as substances that are adsorbed by plastics.

In the case of "topic 37 – NPs aggregation", there seems to be an even recurrence of physical and biological weathering factors, but also material properties. Although it's hard to say for certain what they are referring to, this visualisation helps getting an understanding of particular focuses on this field. From Fig. 4, we also know that molecular weight and crystallinity are material properties that have a high relevance for degradation, especially in biodegradation. The importance

given to them by studies on marine plastics, relative to other material properties reflects this. Interestingly, crystallinity does not appear in biodegradation studies as often as molecular weight and that could be because the latter would appear as a product of degradation as well – so it would be more recurrent, than crystallinity. Also, the reduction of the molecular weight or the formation of carbonyl groups are the main ways to verify degradation, so it was expected to see these terms being frequently mentioned. Interestingly, this is not the case for the latter.

The term "bond energy", which is the main property to be considered in the degradation of polymers, doesn't appear in any of the topics. Although it could be addressed indirectly by
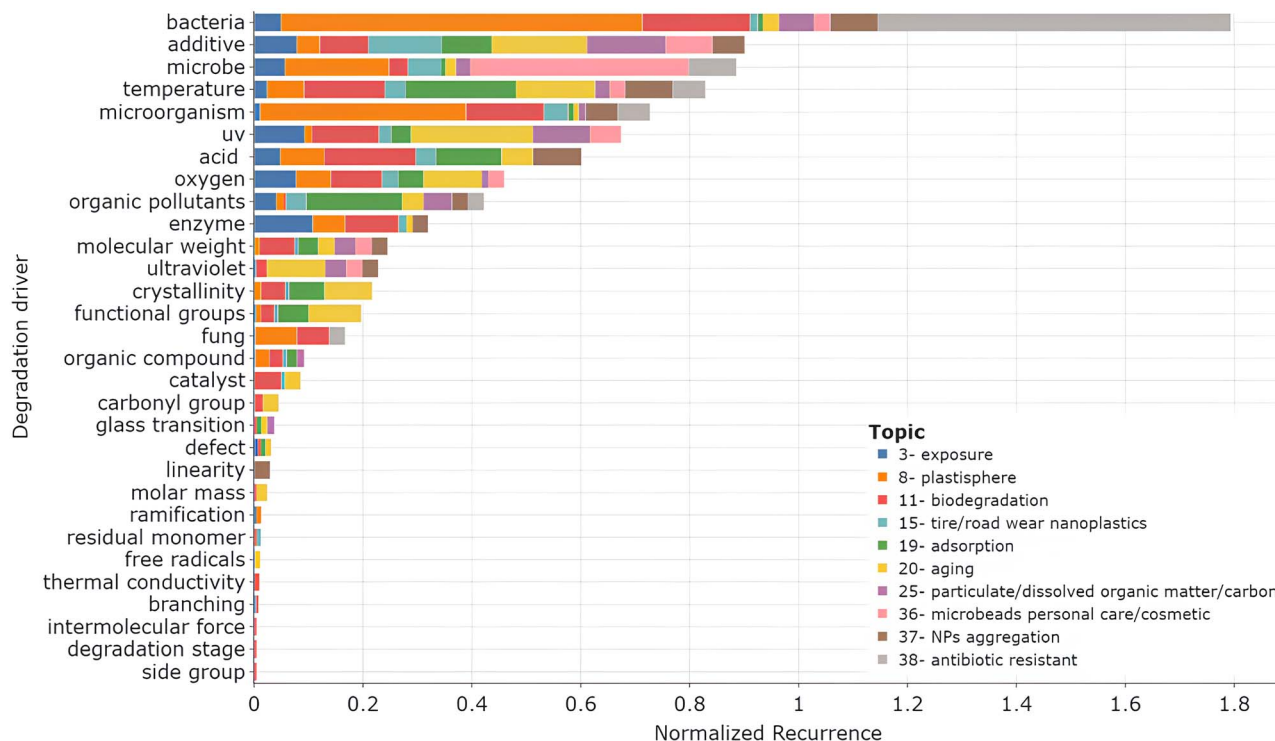
Fig. 7 Normalized recurrence of drivers by topic (colour) for top 10 topics.

other material properties, this doesn't seem to be the case. The recurrences of "free radical" and "chromophore" follow the same tendency. Other important properties such as chain mobility, which can be addressed indirectly by linearity, ramifications or the glass transition temperature are rarely present. It is true that crystallinity also plays a role in chain mobility. However, it may often be addressed as a factor of increased density and reduced intermolecular space, which hinders degradation, especially by biological drivers. Nevertheless, this highlights the fact that a more in-depth focus on polymer properties and degradation mechanisms is still lacking.

Conclusions

ML and NLP tools boosted the process of literature screening and topic understanding. Nevertheless, they can't be used automatically without human intervention. Removing and merging topics as well as constructing a vocabulary to search demands expert knowledge. While there is a lot more they can accomplish and future developments promise greater innovations, artificial intelligence does not supersede human intelligence. The usefulness of these tools demonstrates that they in fact expand our capacity to grasp and build knowledge.

The finding that biological degradation is more prevalent in the marine plastics literature, might be because ecologists are commonly looking at microorganisms since they are a major component of the marine environment. In spite of that, there is a lack of previous understanding of this kind of degradation. Because polymer scientists are mostly focused on designing materials for applications, they study degradation mainly under

processing and use conditions. Since biological degradation is minimal under these conditions, it only appears as a topic of great interest in the marine plastics community, which is concerned about the consequences of the end-of-life of plastic products.

We advocate for the complementarity of polymer science and marine plastics research and praise the efforts of publishers to bring different disciplines together, such as in this very journal. Our results indicate that there is a big opportunity for researchers working in environmental plastics to contribute to polymer science, considering the gap of biodegradation in polymer science literature, as a study on one may turn out to be influential on the other. Yet, for it to have a greater impact across fields, it must be aware of the vocabulary and significance of at least the most important polymer properties. In this work, we intended to give an overview of degradation processes and polymer properties that many marine scientists may not be fully aware of. Although not exhaustive, this basic dictionary aims to enrich the vocabulary and facilitate the connection between terms that may seem unrelated at first for researchers not trained as polymer scientists. We understand that without a thorough consideration of the polymer properties that arise from their composition and structure, the behaviour of plastic particles cannot be fully understood. Therefore, their fate and accumulation characteristics can be inappropriately defined.

Author contributions

Henrique de Medeiros Back: conceptualization, data curation, software, formal analysis, methodology, writing – original draft;

Camila Andreussi: investigation; Daphiny Pottmaier: conceptualization, methodology, writing – review & editing; Orestes Estevam Alarcon: conceptualization, supervision.

Conflicts of interest

There are no conflicts of interest to declare.

Notes and references

1 L. Simon-Sánchez, M. Grelaud, C. Lorenz, J. Garcia-Orellana, A. Vianello, F. Liu, J. Vollertsen and P. Ziveri, *Environ. Sci. Technol.*, 2022, **56**, 16780–16788.

2 J. Zalasiewicz, C. N. Waters, J. A. Ivar do Sul, P. L. Corcoran, A. D. Barnosky, A. Cearreta, M. Edgeworth, A. Gałuszka, C. Jeandel, R. Leinfelder, J. R. McNeill, W. Steffen, C. Summerhayes, M. Wagreich, M. Williams, A. P. Wolfe and Y. Yonan, *Anthropocene*, 2016, **13**, 4–17.

3 R. C. Thompson, S. H. Swan, C. J. Moore and F. S. vom Saal, *Philos. Trans. R. Soc., B*, 2009, **364**, 1973–1976.

4 GESAMP, *Guidelines for the Monitoring and Assessment of Plastic Litter in the Ocean*, ed. P. J. Kershaw, A. Turra and F. Galgani, Rep. Stud. GESAMP No. 99, 2019, p. 130.

5 K. L. Law, S. Morét-Ferguson, N. A. Maximenko, G. Proskurowski, E. E. Peacock, J. Hafner and C. M. Reddy, *Science*, 2010, **329**(5996), 1185–1188.

6 A. Cózar, M. Sanz-Martín, E. Martí, J. I. González-Gordillo, B. Ubeda, J. Á. gálvez, X. Irigoien and C. M. Duarte, *PLoS One*, 2015, **10**(4), e0121762.

7 J. A. Ivar do Sul, A. S. Tagg and M. Labrenz, *Scientometrics*, 2018, **117**, 2145–2157.

8 A. Cohan, S. Feldman, I. Beltagy, D. Downey and D. S. Weld, arXiv, 2020, preprint, arXiv:2004.07180, DOI: **10.48550/arXiv.2004.07180**.

9 S. S. Stivala and L. Reich, *Polym. Eng. Sci.*, 1980, **20**(10), 654–661.

10 P. R. Sundararajan, Chain Structures, *Physical Properties of Polymers Handbook*, ed. J. E. Mark, Springer, New York, NY, 2nd edn, 2007, pp. 3–24.

11 C. K. Jayasuriya and J. K. Premachandra, *Physical Properties of Polymers Handbook*, ed. J. E. Mark, Springer New York, New York, NY, 2007, pp. 25–40.

12 L. S. Reich and S. Stivala, *J. Pharm. Sci.*, 1970, **59**(1), 131–132.

13 Y.-R. Luo, *Comprehensive Handbook of Chemical Bond Energies*, CRC Press, 2007.

14 J. E. Huheey, E. A. Keiter and R. L. Keiter, *Inorganic Chemistry: Principles of Structure and Reactivity*, Pearson Education, 4th edn, 1997.

15 J. B. Huang, G. S. Zeng, X. S. Li, X. C. Cheng and H. Tong, *IOP Conf. Ser.: Earth Environ. Sci.*, 2018, **167**, 012029.

16 A. A. Askadskii, M. Popova, T. Matseevich and E. Kurskaya, *Adv. Mater. Res.*, 2014, **864–867**, 751–754.

17 W. L. Hawkins, W. Matreyek and F. H. Winslow, *J. Polym. Sci.*, 1959, **41**, 1–11.

18 G. Ayoub, F. Zaïri, C. Fréderix, J. M. Gloaguen, M. Naït-Abdelaziz, R. Seguela and J. M. Lefebvre, *Int. J. Plast.*, 2011, **27**(4), 492–511.

19 N. P. Cheremisinoff, *Condensed Encyclopedia of Polymer Engineering Terms*, 2001, 39–81.

20 X. Zhu, Y. Zhou and D. Yan, *J. Polym. Sci., Part B: Polym. Phys.*, 2011, **49**, 1277–1286.

21 T. P. Lodge and P. C. Hiemenz, in *Polymer Chemistry*, 3rd edn, 2020.

22 E. M. James, *Physical Properties of Polymers Handbook*, Springer, New York, NY, 2nd edn, 2007.

23 D. Feldman, *J. Polym. Environ.*, 2002, **10**, 163–173.

24 V. V. Krongauz, *Thermochim. Acta*, 2010, **503–504**, 70–84.

25 S. Girois, L. Audouin, J. Verdu, P. Delprat and G. Marot, *Polym. Degrad. Stab.*, 1996, **51**, 125–132.

26 P. Gijsman, *e-Polymers*, 2013, **8**(1), 065.

27 S. Leo Madorsky, *Thermal Degradation of Organic Polymers*, John Wiley & Sons, Inc, 1st edn, 1964.

28 V. V. Kholodovych and W. J. Welsh, Thermal-Oxidative Stability and Degradation of Polymers, in *Physical Properties of Polymers Handbook*, ed. E. M. James, Springer New York, NY, 2nd edn, 2007, pp. 927–938.

29 J. Izdebska, *Printing on Polymers: Fundamentals and Applications*, 2016, pp. 353–370.

30 P. Gijsman, J. Hennekens and J. Vincent, *Polym. Degrad. Stab.*, 1993, **42**, 95–105.

31 J. Vohlidal, *Chem. Teach. Int.*, 2021, **3**, 213–220.

32 S. J. Joshi, S. Kant Bhatia, K. University, S. Korea, W.-M. Wu, D. B. Levin, N. Mohanan, Z. Montazer and P. K. Sharma, *Front. Microbiol.*, 2020, **11**, 580709.

33 J. Yuan, J. Ma, Y. Sun, T. Zhou, Y. Zhao and F. Yu, *Sci. Total Environ.*, 2020, **715**, 136968.

34 T. Kijchavengkul and R. Auras, *Polym. Int.*, 2008, **57**, 793–804.

35 G. Kale, T. Kijchavengkul, R. Auras, M. Rubino, S. E. Selke and S. P. Singh, *Macromol Biosci.*, 2007, **7**, 255–277.

36 S. Toze and J. Sidhu, *Encyclopedia of Environmental Health*, 2011, DOI: **10.1016/B978-0-444-52272-6.00372-X**.

37 A.-U.-R. Bacha, I. Nabi, M. Zaheer, W. Jin and L. Yang, *Sci. Total Environ.*, 2023, **858**, 160108.

38 R. Egger and J. Yu, *Front. Sociol.*, 2022, **7**, 886498.

39 E. Callaway, H. Ledford, G. Viglione, T. Watson and A. Witze, *Nature*, 2020, **588**, 550–552.

40 E. Gibney, *Nature*, 2020, **582**, 173–174.