Environmental Science & Technology

pubs.acs.org/est

Article

Polymer Biodegradation in Aquatic Environments: A Machine Learning Model Informed by Meta-Analysis of Structure-Biodegradation Relationships

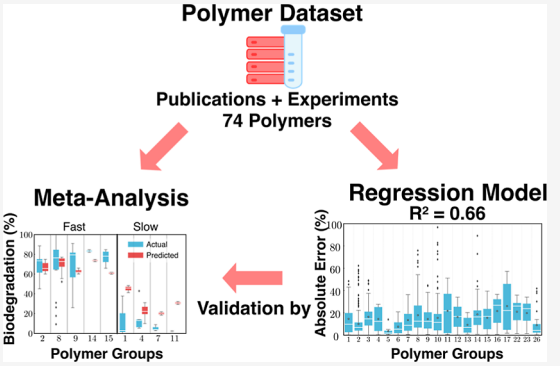
Chengrui Lin and Huichun Zhang*

Cite This: *Environ. Sci. Technol.* 2025, 59, 1253−1263

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Polymers are widely produced and contribute significantly to environmental pollution due to their low recycling rates and persistence in natural environments. Biodegradable polymers, while promising for reducing environmental impact, account for less than 2% of total polymer production. To expand the availability of biodegradable polymers, research has explored structure-biodegradability relationships, yet most studies focus on specific polymers, necessitating further exploration across diverse polymers. This study addresses this gap by curating an extensive aerobic biodegradation data set of 74 polymers and 1779 data points drawn from both published literature and 28 sets of original experiments. We then conducted a meta-analysis to evaluate the effects of experimental conditions, polymer structure, and the combined impact of polymer structure and properties on biodegradation. Next, we developed a machine learning model to predict polymer biodegradation in aquatic environments. The model achieved an R^2_{test} score of 0.66 using Morgan fingerprints, detailed experimental conditions, and thermal decomposition temperature (T_d) as the input descriptors. The model's robustness was supported by a feature importance analysis, revealing that substructure R−O−R in polyethers and polysaccharides positively influenced biodegradation, while molecular weight, T_d, substructure −OC(═O)− in polyesters and polyalkylene carbonates, side chains, and aromatic rings negatively impacted it. Additionally, validation against the meta-analysis findings confirmed that predictions for unseen test sets aligned with established empirical biodegradation knowledge. This study not only expands our understanding across diverse polymers but also offers a valuable tool for designing environmentally friendly polymers.



KEYWORDS: *plastic, polymer, inherent biodegradation, ready biodegradation, OECD 301, mineralization, machine learning*

INTRODUCTION

Polymers are produced on a large scale, with 400.3 million tons of solid polymers (plastics) and 36.3 million tons of liquid polymers produced annually, and are widely used across industries such as packaging, agriculture, and cosmetics.[1−4] However, up to 94% of these polymers end up as environmental pollutants, contributing to issues like microplastics.[5] Conventional polymers exhibit high stability and resistance to microbial degradation in the environment, while biodegradable alternatives, which can significantly reduce environmental lifetimes from hundreds of years to less than a year or even just a few days, have been introduced in various industries to reduce environmental impact.[6−8] Despite this, biodegradable polymers constitute less than 2% of total production, indicating a need for research to expand their availability.[9]

Biodegradation of polymers is categorized into primary and ultimate biodegradation (mineralization).[10] Primary biodegradation breaks polymers into smaller molecules that may persist in the environment, while ultimate biodegradation fully converts them into CO_2, H_2O, CH_4, and biomass under aerobic and anaerobic conditions, ensuring complete biodegradability.[10] Aerobic biodegradation is particularly significant due to oxygen's prevalence. Enhancing polymer biodegradability typically involves structural modifications, blending biodegradable polymers, and additives.[11−13] Because polymer biodegradability depends on hydrolyzable bonds in their structures, understanding the structure-biodegradation relationship (SBR) is especially attractive for assessing environmental risks, improving existing polymers, and designing new, more degradable ones.[10,13]

Around 2000 studies on polymer biodegradation, indexed in Web of Science, cover various scales from laboratory to field tests and use diverse analytical methods. Techniques include polymer-specific methods like gas chromatography, high-

Received: October 18, 2024
Revised: December 21, 2024
Accepted: December 23, 2024
Published: January 7, 2025

https://doi.org/10.1021/acs.est.4c11282
Environ. Sci. Technol. 2025, 59, 1253−1263

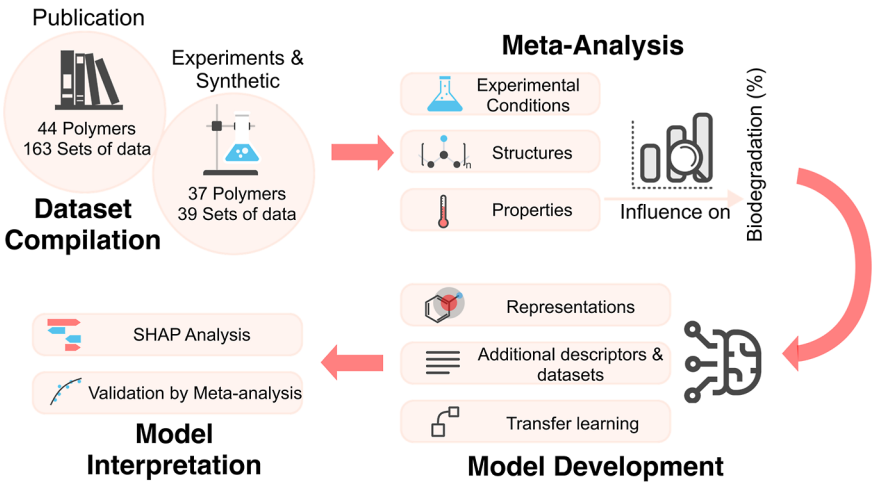


Figure 1. Workflow of the study. (1) Data compilation from the literature, experiments, and synthetic data. This included 163 sets of reported time series data on 44 polymers, as well as 37 polymers and 39 sets of experimental and synthetic data. (2) Meta-analysis to assess the impact of experimental conditions, polymer structures, and polymer properties on biodegradation. (3) ML modeling through performance comparisons of various polymer representations while also exploring alternative approaches such as transfer learning, the incorporation of molecular descriptors, and the integration of experimental and synthetic data sets. (4) Model interpretation using SHAP analysis and by comparison with the meta-analysis findings.

performance liquid chromatography, mass spectrometry, and ^{14}C-labeling, as well as nonspecific approaches such as oxygen consumption and CO_2 evolution measurements.[10,14−16] Our latest work also introduced a high-throughput closed bottle test method with fiber-optic dissolved oxygen measurements to increase efficiency.[17] Other studies have developed automated monitoring systems to reduce labor.[18] However, these methods, typically requiring at least 28 days, are time-consuming and involve specialized equipment.

Alternatively, SBRs have been widely explored to assess polymer biodegradability based on molecular structures. Factors like carbon chain length, aromatic components, and functional groups—such as side chains, ether bonds, and double bonds—have been evaluated, primarily for polyesters, from polyalkylene dicarboxylates (including diols and dicarboxylates (diacids), referred to as diol-diacid polyesters hereafter) to more complex polyesters (polyoxabicyclates and polyisosorbide-*co*-diol oxalates) and to polyester-amides.[11,19−24] Also, studies have examined SBRs for polysaccharides, polyaspartic acid, and polyurethanes.[25−27] Details on these SBRs are summarized in Table S1 in the Supporting Information. Despite this progress, these relationships are generally limited to about ten individual polymer types, with each relationship typically covering only one type of polymer (generally fewer than ten polymers per type). Additionally, current review papers primarily provide summaries within the same polymer type.[28,29] Thus, these relationships have not yet been extended to diverse polymers across different types. To fill this gap, more advanced methods such as machine learning (ML) modeling are needed to evaluate polymer biodegradation, enabling more efficient, broad-spectrum screening and facilitating the design of biodegradable polymers. Surprisingly, only one recent study reported ML modeling on polymer biodegradation, focusing on classifying primary biodegradation (biodegradable or not) of 642 water-insoluble polyesters and polycarbonates under fixed conditions with one bacterial strain.[30]

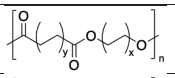
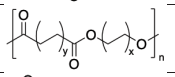
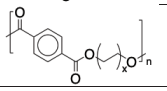
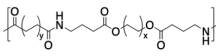
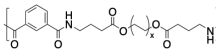
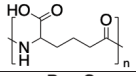
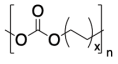
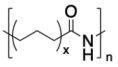
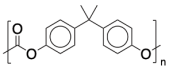
Another major challenge to modeling polymer biodegradation is data scarcity as the majority of studies focus on easily biodegradable polymers like polyethylene glycol (PEG), starch, poly(vinyl alcohol) (PVA), and polyhydroxybutyrate (PHB), limiting structural diversity. Existing biodegradation data are also under different experimental conditions such as inoculum type and amount, substance concentration, temperature, mineral medium concentration, biodegradation end points, and analytical methods. While some studies follow standard guidelines (e.g., OECD and ISO), others modify or do not adhere to them, such as using nonstandard inocula or isolated bacterial strains, complicating comparisons across studies.[31−33]

While there are no regression models for polymer biodegradation, numerous ML models using quantitative structure−property relationships (QSPRs) exist to predict over 20 polymer properties, such as density, melting temperature (T_m), and glass transition temperature (T_g).[34,35] Additionally, rank-based ML models have been developed to assess the abiotic degradability (not biodegradability) ranking of 39 polymers in artificial seawater.[36] Also, our previous work developed an aerobic biodegradation model predicting biodegradation extent based on data for 6000 small molecules and an anaerobic model predicting half-life based on 206 small molecules under varying conditions.[37,38]

Given the limitations of current polymer biodegradation models and the success of polymer property and small-molecule biodegradation prediction models, we aim to develop a regression model to predict polymer aerobic biodegradation in aquatic environments. This model accounts for the ultimate biodegradation of various types of polymers under different conditions over time. To tackle data scarcity, we compiled data sets from both the literature and our experiments, conducting 28 biodegradation tests on 26 different polymers, following the OECD 301B and 301D guidelines to ensure consistency in experimental conditions. To further expand the data set, 11 polymers classified as nonbiodegradable but lacking experimental data were added to our synthetic data set. We then conducted meta-analysis to assess the impact of experimental conditions, polymer structures, and polymer properties on biodegradation. Next, various molecular descriptors were used to represent polymer repeat unit structures, and ML algorithms

https://doi.org/10.1021/acs.est.4c11282
Environ. Sci. Technol. 2025, 59, 1253−1263

Table 1. Summary of Polymers[a]

Polymer Group	Group No.	Structure	Description	Findings in Meta-analysis	P	E	S	Total
Short chain polyester	1		x=1 with or without side chain	Shorter chain higher rigidity and lower biodegradability	1	2	0	2
Longer chain polyester	2		x=2 with side chain, x=2 different side chain length copolymers, or x=5	Longer chain reduced rigidity, higher biodegradability	4	0	0	4
Shorter diacid polyester	3		y=2 with single bond, double bond or side chain; x= 2 or 4	Diacid higher rigidity and lower microorganism populations	4	1	0	4
Longer diacid polyester	4		x = 4; y = 2 or 4		3	2	0	4
Aromatic polyester	5		x= 2 or 3; 1 ring or 2 rings; copolymer of aliphatic-co-aromatic		1	0	4	5
Aliphatic polyester-amide	6		y = 2 or 4; x = 2 or 4		5	0	0	5
Aromatic polyester-amide	7		x = 2, 3 or 4	Aromatic rings lower biodegradability	3	0	0	3
Polyether	8		x = 2, 2 with side chain, or 3	High solubility and naturally occurring microorganisms	2	2	0	3
Polyamino acid	9		2 polyaspartic acid isomer and one polyglutamic acid	Compose of amino acids which are degraded by naturally occurring microorganisms	2	1	0	3
Acrylic-based polymers	10		$R_2 = CH_3$, $R_1 = OCH_3$; R_2 = H, $R_1 = NH_2$, OH; R_2 = OH, R_1 = OH	Stiff C-C backbone and higher T_m, lower biodegradability	5	2	2	8
Polyalkylene Carbonates	11		x = 2, 3, 4		1	3	0	4
Polyamide	12		1, 2 amide bond and different carbon length		2	1	3	6
Polyolefin	13		x = 2 with or without side chain		1	0	1	2
Polysaccharides	14		Natural polymers	Naturally occurred, higher biodegradability	6	2	0	7
Poly(vinyls)	15-20		Different functional group in side chain	PVA: Water soluble and microorganisms adapted	2	3	1	5
Others	21-24				2	2	0	4
High performance polymer with aromatic structures	25				0	5	0	5
Total					44	26	11	74

[a]*P*: number of polymers in the publication data set; *E*: number of polymers in the experimental data set; *S*: number of polymers in the synthetic data set. Note: 7 polymers present both in publication and experimental data sets.

were screened. Additionally, polymer properties T_m and thermal decomposing temperature (T_d) predicted by Polymer Genome were included as descriptors.[35] Finally, the model was interpreted by Shapley additive explanations (SHAP) analysis[39] and validated by comparison with our meta-analysis findings.

METHODS

Data Set Compilation. The polymer biodegradation data were compiled from three sources: journal publications, experimental data, and synthetic data (Figure 1). For journal publications, relevant literature containing biodegradation data was identified through an initial screening process using both general keywords (e.g., "polymer biodegradation") and specific keywords combining polymer names and guideline names (Table S2) via Google Scholar. The data was subsequently extracted, screened, and cleaned following the procedures in Text S1.1. Table S3 summarizes the polymer abbreviations mentioned in this study.

For the experimental data, we conducted biodegradation tests on 26 commercially available polymers that lacked biodegradation data, showed significant variance in published results, or were tested under nonstandard guidelines (details of experiments in Text S1.2 and Table S4). The synthetic data included 11 polymers reported as nonbiodegradable but lacking biodegradation tests. These were generated by assigning 0% biodegradation at day 28 under OECD

https://doi.org/10.1021/acs.est.4c11282
Environ. Sci. Technol. 2025, 59, 1253−1263

301D—the most stringent guideline, where polymers are least likely to pass—to increase data diversity.[40]

The model input features included polymer repeat unit structures, molecular weight (M_w), experimental conditions predefined by guidelines, and reaction time, and the output feature was biodegradation percentage (details in Table S5). Please refer to the Excel file in the Supporting Information for the data set.

Polymer and Guideline Representations. Polymers were represented by their repeating units in SMILES notation, with termini marked by asterisks (*), then converted into three molecular fingerprints: MACCS, Morgan (radius = 2, bits = 2048), and count-based Morgan (radius = 2, bits = 2048; referred to as Morgan with counts hereafter).[41] Additionally, two embeddings were used: molecular embedding (ME) and TransPolymer embedding.[42,43] The TransPolymer embedding was extracted from TransPolymer repretrained by adding our polymers to the original corpus to tailor it for our task (details in Text S1.5).

Standard guidelines were initially encoded categorically into guidelines and principles, following our previous work.[37] However, categorical encoding treats each guideline independently, failing to capture the interrelationships. To improve this, guidelines were also detailed by experimental conditions, including substance concentration levels, inoculum levels, mineral medium concentrations, measurement methods, and acclimated conditions (Table S5).

Polymer Groups for Meta-Analysis and Modeling. Different grouping strategies were employed for meta-analysis and for modeling. For meta-analysis, polymers were grouped based on the unique functional groups present in their structures to ensure group similarity (Table 1). For modeling, however, some groups were too small to achieve a similar polymer group distribution across the training and test sets. To address this, certain groups were combined to increase the polymer count in each group while preserving similarity. For example, polyesters were not divided into short-chain and long-chain categories.

Meta-Analysis. Data obtained from kinetic curve fitting using the Hill sigmoid model, which has previously shown a good fit for biodegradation data (details in Text S1.3), was used in meta-analysis to impute missing data points at day 28 as not every set of data included measurements for this specific day.[44] Day 28 was chosen as it aligns with the guideline end point and ensures test duration consistency.

The analysis began by evaluating how conditions—such as substance concentration, inoculum level, and measurement method—affected biodegradation. Only ready guidelines were analyzed due to limited inherent guidelines. Subsequently, data with similar conditions were combined to assess the influence of polymer groups (representing different structures) on biodegradation, minimizing the external factor impacts. Statistical differences between guidelines or between polymer groups were examined using the Tukey HSD test. Furthermore, within certain groups, the effects of polymer substructures, properties predicted by Polymer Genome that reportedly potentially influence biodegradation (Text S1.4), and M_w on biodegradation were examined.[35,45,46]

Model Development. Regression models were developed to predict biodegradation percentages, initially screening 11 algorithms: decision tree regressor, elastic net, gradient boosting regressor, K-neighbors regressor (KNN), LASSO, linear regression, random forest (RF) regressor, read-across, ridge, support vector regression, and XGBoost regressor (XGB).[47] These models were trained using MACCS, selected for its strong performance in small-molecule biodegradation predictions.[37] Model training and tuning followed nested stratified group cross-validation (NSGCV, Text S1.6), where each set of time series data was grouped to ensure that it appeared only in training or testing to prevent data leakage. Polymer groups were stratified to ensure diversity and similarity of groups between training and test sets, thus enhancing robustness. Model performance was evaluated by R^2 and root-mean-square-error (RMSE) values. Given the absence of benchmark data sets and models for polymer biodegradation, KNN and read-across were employed as baselines. Their nonparametric nature avoids parameter fitting, relying on minimal assumptions. Additionally, predictions from the small-molecule biodegradation model, when applied to polymers, were included as another baseline.[37]

Hyperparameters of the best performance algorithms were tuned using the tree-structured Parzen estimator.[48] Different polymer representations were compared to assess their effectiveness. The model was also trained using an expanded data set using kinetic fitting (Text S1.3) or adding experimental and synthetic data to increase the diversity of polymers and conditions. This additional data set was used only in the training set to maintain a consistent test set for model evaluation. Additional polymer properties, predicted by Polymer Genome, were included as features to provide information beyond the repeating units.[49]

To assess whether the large, small-molecule biodegradation data set could improve polymer predictions, transfer learning was employed using a model pretrained on the small-molecule biodegradation data.[37] Additionally, the potential of language models was explored by fine-tuning the TransPolymer model on the biodegradation data set.[43] Since both approaches used neural network architectures, a neural network model without transfer learning was also trained on the polymer biodegradation data to serve as a baseline model. Details of the above approaches are in Text S1.7. Classification models were also developed to predict biodegradation levels (details in Text S1.8).

Application Domain (AD). The absolute prediction errors were computed for all data points following the NSGCV approach (Text S1.6), where each data point appeared in the training set for some splits of cross-validations and in the test set for other splits. Consequently, training set errors for each data point were calculated as the average prediction errors for certain splits, where the data point was only in the training set. A similar approach was applied for test set errors. This enabled the calculation of prediction errors for both the training and test sets across the entire data set. For each polymer group, if no statistically significant difference in prediction errors was observed between the training and test sets, then the group is considered within the domain. In such cases, the interquartile range (IQR) can define the application domain (AD) for that group.

Model Interpretation. To understand how the models made predictions, SHAP analysis was utilized to assess feature importance by calculating the SHAP and their average absolute values for each feature.[39] These values indicated how much each feature influenced the biodegradation predictions, positively or negatively.

Further validation involved domain-based interpretation, examining whether the model's predictions were consistent

https://doi.org/10.1021/acs.est.4c11282
Environ. Sci. Technol. 2025, 59, 1253−1263

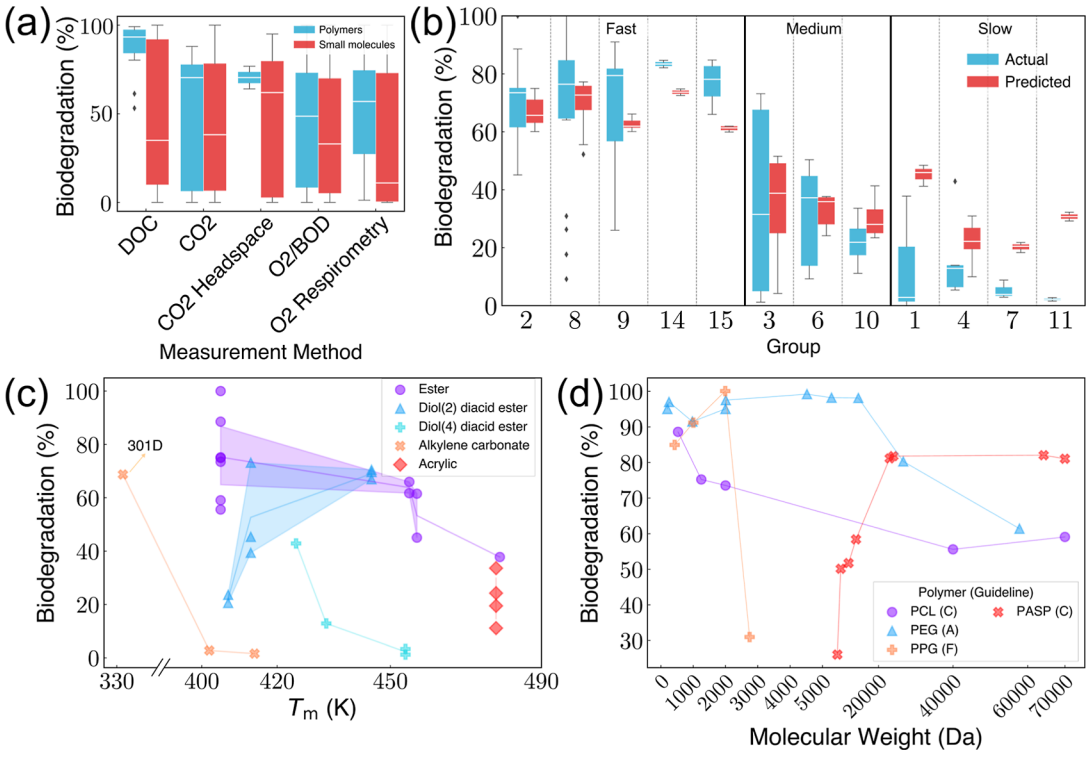


Figure 2. (a) Influence of measurement methods on biodegradation for the polymer data set (blue) and small-molecule data set (red). (b) Influence of polymer groups on biodegradation, showing actual (blue) and predicted (red) values from the best-performing model under OECD 301B, C, and F, segmented by degradation rate (fast, medium, and slow). (c) Effect of melting temperature (T_m) on biodegradation for different polymer groups under OECD 301C and F, except for one polyalkylene carbonate tested under 301D. (d) Effect of molecular weight on biodegradation for polycaprolactone (PCL), polyethylene glycol (PEG), polypropylene glycol (PPG), and polyaspartic acid (PASP). The letters following each polymer name in the legend indicate the OECD 301 guidelines under which these polymers were tested (more details in Figure S11).

with established empirical biodegradation knowledge. The meta-analysis findings were employed to validate the model outputs. Predictions were generated by averaging the values on the test set using the NSGCV strategy (analogous to calculating test errors for AD but only using predicted values here), ensuring that the data were unseen by the model. These predictions were initially compared to known biodegradation trends across different polymer groups, followed by an analysis within specific polymer groups. Intragroup trends were assessed using Pearson correlation coefficients (r) to evaluate the agreement between predicted and observed biodegradation values.

RESULTS AND DISCUSSION

Data Set Summary. First, biodegradation data for 44 different polymers were collected from 40 screened publications (details of the polymers in Table 1).[50,51] This publication data set includes 163 individual sets of biodegradation data. Each set represents a biodegradation kinetic curve characterized by a fixed polymer structure, M_w, and experimental conditions. When discrete time points within these kinetic curves are considered, the data set comprises 1515 distinct data points. Second, the experimental data set includes 26 polymers, 7 of which are also present in the publication data set (details in Table S4), contributing 28 sets of data (253 data points). Finally, the synthetic data set comprises 11 polymers with 11 sets of data (11 data points). Overall, the combined data set spans 74 polymers, 202 sets of biodegradation data, and 1779 individual data points.

Among the 202 sets of data, 146 sets contain fewer than 10 data points, 46 sets include between 10 and 20 data points, and 10 sets include more than 20 data points (Figure S4a). Regarding test duration, 116 sets have a maximum duration of 20−40 days, 69 sets exceed 40 days, and 17 sets are under 20 days (Figure S4b). For the extent of biodegradation, approximately half of the data (103 sets) report values below 60%, while the remaining sets exceed 60% (Figure S4c). In terms of M_w, 29 sets lack M_w data, 66 sets have M_w between 1000 and 10,000 Da, 42 sets range from 10,000 to 100,000 Da, and the remaining sets are higher than 100,000 Da (Figure S4d). For guidelines, OECD 301B is the most used (57 sets), followed by 301F (41 sets), 301A (35 sets), and 301D (32 sets) (Figure S4e). Additionally, most experiments (177 sets) used nonacclimated sludge, while 25 used acclimated sludge.

The 74 polymers were grouped into 25 groups. Detailed polymer groups are provided in Table 1 and explained in Text S2.1. Since some groups were combined, the polymer groups for NSGCV in ML were consolidated into 14 groups, as summarized in Table S6.

Meta-Analysis. A meta-analysis was conducted to analyze factors influencing biodegradation. Of the 202 sets of data, 188 adhered to the ready biodegradability guidelines (OECD 301 series and 310), with 140 having biodegradation results available at 28 days. Polymer groups with fewer than two sets of data were excluded due to a lack of representativeness, leaving 130 data sets. Only biodegradation data obtained under nonacclimated inoculum conditions were considered, ultimately yielding 121 data sets for analysis.

https://doi.org/10.1021/acs.est.4c11282
Environ. Sci. Technol. 2025, 59, 1253−1263

Experimental Conditions. To evaluate how experimental conditions influence biodegradation, we first examined the impact of biodegradation quantification methods (Figure 2a), where the dissolved organic carbon (DOC) method showed higher biodegradation than the other measurement methods, likely due to DOC removal by adsorption on biomass.[52] Then, no statistically significant differences were observed among the other guidelines (not including DOC-based 301A) or for other conditions, such as the substance concentration, inoculum concentration, and mineral medium ratio (Figures 2a and S5a−d). Similar trends were observed between the small-molecule and polymer data sets, with the latter showing more pronounced trends, likely due to its narrower range of data.

Comparison among Polymer Groups. To evaluate the influence of polymer groups, we compared the biodegradation extent among different groups. To minimize the impact of other factors, data from OECD B, C, and F were combined due to their similar biodegradation profiles. All of the comparisons below are based on these guidelines unless specified otherwise. After comparing biodegradation across all polymer groups, we identified group pairs with statistical differences based on Tukey HSD and categorized them into fast, medium, and slow groups (Figure 2b): all the fast groups consistently exhibited higher biodegradation than all the slow groups; the medium groups were degraded more slowly than some fast groups, but their biodegradation was not statistically different from the slow groups.

Among the fast groups, longer-chain polyesters (Group 2) demonstrated lower rigidity due to their extended chains. Polyethers (Group 8) exhibited good solubility. Polyamino acids (Group 9), composed of amino acids and polysaccharides (Group 14), are natural polymers. The biodegradability of Groups 8, 9, and 14 is associated with the presence of naturally occurring microorganisms.[53−55] In addition, poly(vinyl alcohol)s (Group 15) are water soluble, and their widespread use in detergent capsules may have led to microbial adaptation in sewage systems.[56]

Among the medium biodegradation groups, shorter-diacid polyesters (Group 3) degraded more slowly, likely due to their rigid diacid ester structures and the limited population of microorganisms capable of degrading them.[19,57] Although longer-chain diacid polyesters (Group 4) exhibit lower crystallinity and lower T_m, which theoretically should increase biodegradability compared to Group 3, this was not observed here, likely due to the wide range of biodegradability within Group 3.[19,45] Polyester-amides (Group 6), which contain both diacid ester and amide groups, showed reduced biodegradability. Acrylic-based polymers (Group 10), with their stiff C−C backbones and higher T_m, also exhibited lower biodegradability.[58]

The slow biodegradation groups included short-chain polyesters (Group 1), which have a higher degree of crystallinity, leading to slower degradation.[45] For aromatic polyesteramides (Group 7), the presence of aromatic groups significantly reduced biodegradability. The exact reason for slower biodegradation for polyalkylene carbonates (Group 11) is unclear due to limited biodegradation studies. More discussion of the above groups is in Text S2.2.

Effect of T_m and Substructure on Biodegradation. To evaluate the influence of polymer properties and specific substructures on biodegradation, we compared biodegradation within polymer groups under similar conditions, focusing only on OECD 301C and F, which utilize the same inocula, substance concentration, and measurement methods. T_m was used here because (1) among all the potential properties related to biodegradation (Table S7), only T_g and T_m showed correlation, and (2) T_g and T_m are interrelated, and T_m is more commonly employed to explain observed biodegradability, with a higher T_m indicating less biodegradability.[59]

Initially, we examined the relationship between T_m and biodegradation across four polymer types: polyesters (Groups 1 and 2), diol-diacid polyesters (Groups 3 and 4), polyalkylene carbonates (Group 11), and acrylic-base polymers (Group 10) (Figure 2c). Within each type of polymer, an increase in T_m corresponded to a general decrease in biodegradation except for 2-carbon diol-diacid polyesters (with a fixed diol carbon chain length of four and variable diacid lengths in Groups 3 and 4). Although the melting process is not a prerequisite for biodegradation, T_m is indicative of polymer substructure and chain flexibility, which affects enzymatic accessibility.[59]

To further elucidate the substructural effects on biodegradation, separate analyses were conducted for each polymer group. For the three types of polymers with increased biodegradability as T_m increases: for 4-carbon diol-diacid polyesters (Groups 3 and 4) (Figure S7a) and polyesters (Groups 1 and 2) (Figure S7c), shorter chain lengths in diacid or in the backbone increased rigidity, as reflected by higher T_m, which in turn reduces biodegradability, though side chains may also contribute to lower biodegradation in the latter polymer, consistent with prior reports.[11,19,45,60] In contrast, for polyalkylene carbonates (Group 11), no correlation between chain length and biodegradation was observed (Figure S7d).

For 2-carbon diol-diacid polyesters (Groups 3 and 4), where higher T_m aligns with increased biodegradability, the diacid with the lowest T_m, which contained side chains, exhibited the lowest biodegradability (Figure S7b), potentially due to decreased chain flexibility and reduced enzyme access.[20] This suggests that T_m is not the sole factor influencing biodegradation; other polymer features also play significant roles, which may explain the absence of observable patterns in other polymers. In polyester amides (Groups 6 and 7), the inclusion of aromatic rings in the diacid resulted in increased T_m and decreased biodegradability (Figure S8a). Biodegradation decreased as the diacid length increased under a fixed diol length (Figure S9b) or as the diol chain length increased under a fixed diacid length (Figure S9c).

Despite the above trends, the relationship between carbon chain length and biodegradability is not significant after combining polyesters (Groups 1 and 2), polyethers (Group 8), and polyalkylene carbonates (Group 11) (Figure S10a). For acrylic-based polymers and polyvinyls (Groups 10 and 15−20), polymers with the −OH functional group showed higher biodegradability (Figure S10b).

A detailed discussion of all of these polymer groups is provided in Text S2.3.

Effect of M_w on Biodegradation. The effect of M_w on biodegradation was investigated (Figures 2d and S11). For PCL under 301C and PEG (Figures 2d and S11a, b), biodegradability decreased as M_w increased. This trend was not observed for PCL under 301F (Figure S11a). For PPG (Figures 2d and S11c), biodegradability increased with M_w up to 2000 as lower M_ws with shorter chains exhibited greater rigidity.[61] However, biodegradation declined once M_w exceeded 2000. For poly(aspartic acid) (PASP) under 301C (Figures 2d and S11d), biodegradation increased with M_w, while no obvious trend was observed in 301B (Figure S11d).

https://doi.org/10.1021/acs.est.4c11282
Environ. Sci. Technol. 2025, 59, 1253−1263

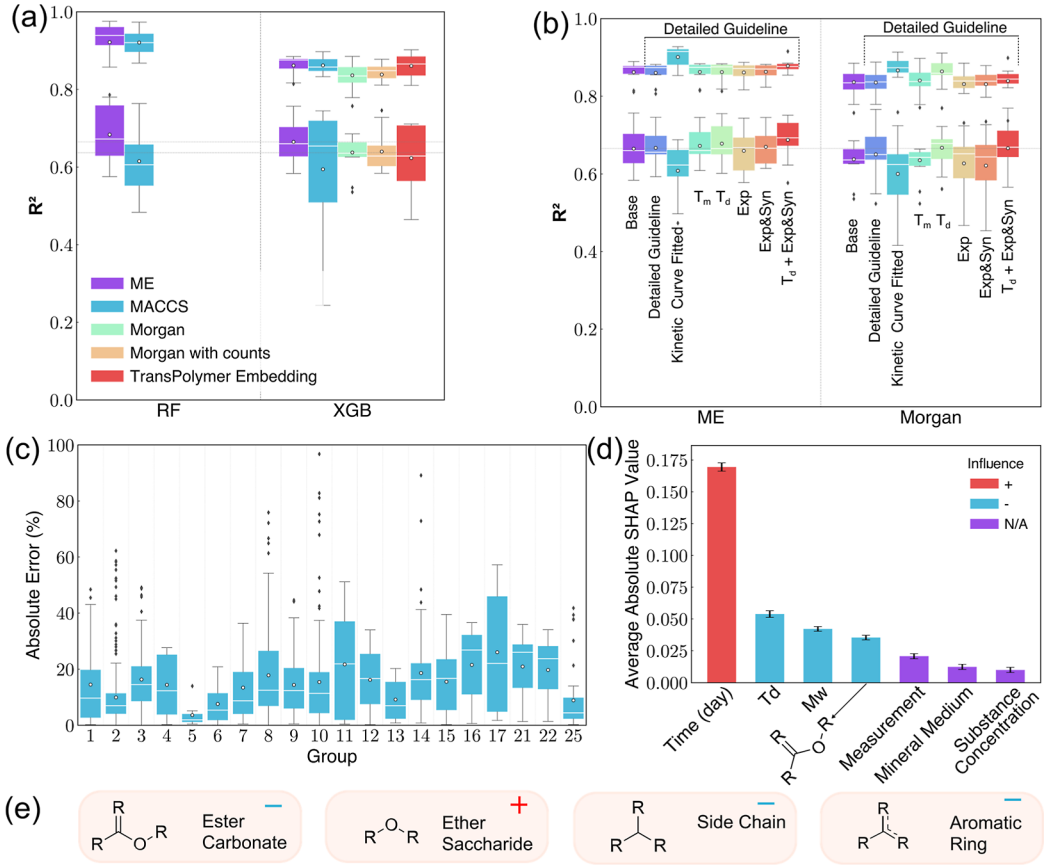


Figure 3. Comparison of R^2 for (a) different polymer representations using RF and XGB models and (b) ME and Morgan representations across various modeling approaches in addition to the base model (categorical guideline representation), including kinetic curve fitting and combinations of experimental (Exp) and synthetic (Exp&Syn) data sets with additional descriptors (T_m and T_d). Note that all models used detailed guidelines except for the base model. Upper boxes indicate training scores, while lower boxes display test scores. (c) AD for polymer groups, showing the test set prediction errors for different polymer groups. New query polymers are expected to have similar prediction errors in IQR if they belong to an existing group. (d) Feature importance by SHAP values for the best regression model (Morgan with detailed guidelines, T_d, and experimental and synthetic data set). Positive influences are shown in red, negative influences in blue, and categorical features with undetermined influence in purple. (e) Four important substructures were identified from the top 18 important features across all input variables. The influences of these features are shown using the + and − signs.

In summary, increasing M_w resulted in reduced biodegradability for PCL and PEG, enhanced biodegradability for PASP, and an initial increase followed by a decrease for PPG. Among PASPs with similar M_ws, PASP with a β-structure biodegraded slowly due to an amide-like structure hindering enzymatic attack (Text S2.4).[26,54]

Model Development. *ML Algorithm and Polymer Representation Prescreening.* In the initial prescreening phase, XGB and RF were the top-performing models (Figure S13) and were subsequently evaluated by using both MACCS and ME representations. XGB was selected for further model development due to lower levels of overfitting (Figures 3a and S14a) (for additional prescreening details, see Text S2.5). When compared to the baseline algorithms, KNN and read-across show average R^2_{test} values of 0.34 and 0.33, respectively (Figure S13), while the small-molecule biodegradation model, when applied to polymers, achieves R^2 = 0.17. In contrast, XGB outperforms these baselines, with an R^2_{test} of 0.59. Various polymer representations were compared by using XGB (Figures 3 and S14a). ME outperformed all other representations, achieving an R^2_{test} of 0.67. The strong performance of ME might be attributed to the continuous vector used. Among the fingerprint-based representations, Morgan and Morgan with counts performed better, both achieving an R^2_{test} value of about 0.64. The predefined MACCS representation had the lowest performance, with an R^2_{test} of 0.60. This could be due to MACCS's inability to capture certain substructures, such as variations in carbon chain lengths (e.g., polyamides 4 and 6 share the same MACCS fingerprint).

Since using a single repeat unit may result in the loss of some bonding information as the start and end atoms are not connected, we attempted to link repeat units to form oligomers (Text S2.7). However, all types of repeat unit representations demonstrated comparable performance (Figure S17), indicating that a single repeat unit is sufficient for polymer representation.

For more advanced model architectures, we tested the TransPolymer approach, which included both the use of its embeddings (Figures 3a and S14a) and fine-tuning (Figure S15) and transfer learning based on the small-molecule biodegradation data (Figure S15). Among these, only the transfer learning models pretrained on the small-molecule data and fine-tuned on the polymer data (R^2_{test} of 0.47 for MACCS and 0.58 for ME) outperformed the baseline NN (R^2_{test} of 0.41 for MACCS and 0.37 for ME). Nonetheless, none of the transfer learning strategies surpassed the performance of the

https://doi.org/10.1021/acs.est.4c11282
Environ. Sci. Technol. 2025, 59, 1253−1263

XGB models. Detailed results of these approaches are in Text S2.8.

Based on these findings, ME embedding-based representation and the Morgan fingerprint-based representation were determined to be the most effective and were selected as the base model for all subsequent modeling work.

Additional Descriptors and Data Sets. Additional descriptors and data sets were incorporated into the base model to assess their impact on model performance (Figures 3b and S14b). Initially, the model based on the detailed guideline representations outperformed the base model, with higher R^2_{test} scores, narrower R^2_{test} ranges, and less overfitting, as indicated by the unchanged training scores. The detailed guideline representations encoded the guidelines into three features, highlighting the dependence of biodegradation on experimental conditions, while the categorical guidelines used only one independent feature, masking the effect of the experimental conditions. As a result, all subsequent modeling approaches adopted the detailed guideline representations.

Data augmentation using kinetic curve fitting resulted in a lower R^2_{test}. This was because the augmented data introduced only additional points within each set of time series data without increasing the diversity of conditions or polymer structures. The higher training scores associated with augmentation also suggested potential overfitting.

T_m, which showed a correlation with biodegradability for certain polymers, was added as an additional descriptor. However, the inclusion of T_m did not improve the performance of ME and slightly decreased the performance of Morgan, likely because T_m only correlated with specific polymer types. In contrast, the addition of T_d improved both models' performance: Morgan's R^2_{test} increased from 0.64 to 0.67, and ME's R^2_{test} rose from 0.67 to 0.68 (reasons discussed later).

The inclusion of experimental and synthetic data sets initially reduced performance due to the increased diversity of polymers, potentially leading the model to learn more generalizable patterns. This diversity helped the model become less overfitted. When T_d was added to the expanded data set, the R^2_{test} achieved 0.67 ($R^2_{training}$ 0.84) for Morgan and 0.67 ($R^2_{training}$ 0.88) for ME. This suggests that the expanded data set, combined with relevant descriptors, enhanced the models' predictive accuracy and applicability across a broader range of polymers. Ultimately, the Morgan model, incorporating detailed guidelines, T_d, and the expanded experimental and synthetic data sets, was chosen as the final model due to its interpretability and performance.

Classification models were also developed to determine whether a polymer was biodegradable or to estimate its approximate level of biodegradability without requiring precise measurements of the biodegradation extent. The best regression model approaches were employed in developing the classification models, and the performance of the classification models is provided in Text S2.6. A binary classification model based on a 20% biodegradation threshold achieved an $F1$ score of 0.84 and an accuracy score of 0.85 using the Morgan.

AD FOR POLYMERS AND EXPERIMENTAL CONDITIONS

Initially, we attempted to define the model AD based on similarity among polymers, following a method used in our previous work.[37] However, the prediction errors across all similarity levels were comparable, indicating that AD could not be defined this way (details in Text S2.9 and Figure S19). Consequently, the AD was defined manually by examining the model's prediction errors for each polymer group. Polymer groups with fewer than three data points were excluded due to the lack of representativeness. No statistical difference presented in training and test set errors for all polymer groups (Figure S20), along with the fact that the test set was unseen data, indicates that test errors can reliably reflect the expected errors for new data. If a new query polymer is similar to an existing group, then its prediction error is likely to align with the IQR of test errors for that group, as demonstrated in Figure 3c.

The AD for the experimental conditions was manually defined similarly. AD for each guideline is illustrated in Figure S21. Guidelines 301D and 301E had higher prediction errors than the other guidelines, which generally had prediction errors within the range of 10−20%. No significant correlation was observed between prediction and variables such as M_ws, acclimation, or reaction time (Figure S22). Overall, it is recommended to use condition features within the same range as those in our data set for optimal prediction accuracy.

Model Interpretation. Since nonlinear models were developed using relatively small data sets, it is crucial to assess whether these models have learned relevant and accurate information about biodegradation. If the models demonstrate accurate learning, we can be more confident about their predictions. To evaluate this, model interpretation was carried out using SHAP analysis for feature importance and by comparing the results to the meta-analysis findings.

Model Interpreted by SHAP Analysis. Feature importance was assessed based on the best-performing Morgan model, indicating whether a feature positively or negatively affects biodegradation predictions. Reaction time was identified as the most influential feature (Figure 3d), which aligns with results from the small-molecule models.[37] The second most important feature, T_d, negatively impacted biodegradation. A higher T_d, indicating greater thermal resistance, showed a clear cutoff at 670 K, where values below this threshold had high SHAP values, and values above had low SHAP values (Figure S23c).[62] Thermal decomposition and biodegradation are likely linked as both involve breaking polymers down into smaller molecules, whether by heat or microbial activity. In fact, similar correlations have been observed, though for polymer blends in compost conditions.[63] M_w ranked third in importance (Figure S23d), typically showing a negative influence on biodegradation as larger polymer molecules are harder to degrade. The remaining features had relatively minor influences on the predictions and are discussed in Text S2.10.

Among the 18 most significant features across all input variables, four important substructures were identified. Each of these structures can be represented by different Morgan bits (Figure 3e). For example, the substructure −OC(═O)− found in ester and carbonate bonds corresponds to four Morgan bits, which are ranked fourth, 10th, 11th, and 13th among all important features. This substructure negatively influences biodegradation (Figures 3d and 3e). Despite that ester bonds are reportedly biodegradable, studies may not account for mineralization or be in aquatic environments.[29] Indeed, our meta-analysis for polymer group comparison revealed that certain ester-containing polymers, such as short-chain polyesters and diol-diacid polyesters, reduced the biodegradability. Morgan bits ranked ninth, 12th, and 16th correspond to the R−O−R groups, which are found in

https://doi.org/10.1021/acs.est.4c11282
Environ. Sci. Technol. 2025, 59, 1253−1263

polyethers and polysaccharides. Although the influence of R−O−R substructures on biodegradability has not been reported for polymer, these polymer groups might exhibit high biodegradability due to their interactions with natural organisms,[53,55] consistent with the findings of meta-analysis. Furthermore, the remaining two substructures, side chains and aromatic rings, which have a negative influence, also align with our meta-analysis. Additionally, some structural attributes may be reflected in T_d. Polymers with T_d values exceeding the 670 K threshold generally exhibit high alkyl content or contain aromatic rings (Table S8).

Interactions among the most important features were examined; however, no significant interaction patterns were identified, as discussed in Text S2.10. Additionally, for the ME, whose features lack direct interpretability due to dimension reduction, the potential physical meaning was inferred by examining their correlations with other descriptors (details in Text S2.11).

Interpretation Based on Biodegradation Rules. To further validate the model, we assess whether the model's predictions align with known biodegradation knowledge. First, we applied the model to predict unseen data and assessed whether the predictions followed the reported rules. Note that these studies might not follow standard guidelines or involve aquatic phases due to the limited number of rules reported under standard guidelines. While the model accurately predicted decreased biodegradability with increasing chain length in certain polyamides or T_m in specific diol-diacid polyesters, it struggled with other polyamides and diol-diacid polyesters of similar structural characteristics and with copolymer biodegradation influenced by monomer ratios as these factors were not part of the input data, and experimental conditions varied from standard guidelines (more discussion in Text S2.13).

Given the limited number of studies reporting SBRs and differing conditions between existing studies and ours, we used our meta-analysis findings to validate the best regression model. The goal was to determine whether the model's predictions for the test set were consistent with these findings. First, for different polymer groups, the predicted fast and medium biodegradation rates closely matched experimental values, while the slow biodegradation rates were overpredicted (Figure 2b). This discrepancy may arise from the presence of similar functional groups, such as ester bonds, in both the medium- and slow-degrading polymer groups. This overlap could lead to overpredictions in the slow-degrading category, where the model might make predictions based on functional group similarity.

For the influence of T_m, the model effectively captured the negative impact of both reduced carbon chain length and increased T_m on biodegradation for 4-carbon chain diol-diacid polyesters (r = 0.99, Groups 3 and 4, Figure S7a) and polyesters (r = 0.80, Groups 1 and 2, Figure S7c). Additionally, the model successfully learned the adverse effects of side chains and T_m on biodegradability for 2-carbon chain diol-diacid polyesters (r = 0.90, Groups 1 and 2, Figure S7b). In the case of polyester-amides (Groups 6 and 7) (Figure S9b,c), the model showed strong correlations between biodegradation and diol chain length (r = 0.96) or diacid chain length (r = 1). For polyalkylene carbonates (Group 11) (Figure S7d), however, predictions were not consistent with the real trends, which may be attributed to inoculum variability and nonlinear effects of chain length on biodegradation within this group, where the actual order of biodegradation followed a trend of 3 > 4 > 2 carbon chain lengths.

In terms of trends related to M_w, the model predictions for PCL (r = 0.92, 301C, Figure S11a) and PEG (r = 0.78, 0.92, and 1 for 301A, B, and F, Figure S11b) aligned well with the meta-analysis. For PPG, the model correctly captured the increase-then-decrease trend (r = 0.24, 0.99, and 1 for 301A, B, and F, Figure S11c). In contrast, the increasing trend was not as apparent for PCL and PASP under 301F conditions (Figure S11a,d). This discrepancy may arise because most of the data indicated a decrease in biodegradation with increasing M_w, leading the model to predict accordingly, as reflected in the SHAP values (Figure S23b).

A more detailed discussion is provided in Text S2.14.

Environmental Significance. This study compiles an extensive data set of polymer biodegradation data in aquatic environments under standard guidelines from both published sources and original experiments. Through meta-analysis, the effects of various features—including experimental conditions, polymer groups, substructures, and polymer properties—on biodegradation were examined. Notably, this research represents the first attempt at developing regression models for polymer ultimate biodegradation. The best-performing model employed the Morgan fingerprint, detailed guidelines, and additional descriptors such as T_d, augmented with experimental and synthetic data sets, achieving an R^2_{test} score of 0.66, comparable to the previous small-molecule biodegradation model, which had an R^2_{test} of 0.54 despite being trained on a much larger data set. The model can be applied to over 20 different polymer groups with a prediction error of less than 20%. The model was interpreted using SHAP analysis and validated by comparing its results to meta-analysis findings. Importantly, the model learned key features and trends identified in the meta-analysis, enhancing the confidence in its use. We also launched the model on a free, user-friendly Web site (https://environ-ai-9d4e000c140e.herokuapp.com/polybio/predict, user guide in Text S2.15) for easy access and usability. Future improvements could include more polymer biodegradation results, capturing more complex features, such as copolymers with different monomer ratios, and incorporating features representing sludge microbial information to enhance the model's predictive power.

This data set represents the first comprehensive compilation on polymer ultimate biodegradation. The meta-analysis not only corroborates existing findings on SBRs but also provides new insights into how biodegradability varies across different polymer groups. By validating the model against these findings, we ensured its robustness, making it a reliable tool for predicting polymer biodegradability. This model holds significant potential for designing environmentally friendly polymers, aiding material scientists, environmental researchers, and sustainability-focused industries in making informed decisions. The findings also support environmental agencies and policymakers in assessing and improving polymer biodegradability, contributing to sustainable solutions for polymer waste challenges.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.4c11282.

https://doi.org/10.1021/acs.est.4c11282
Environ. Sci. Technol. 2025, 59, 1253−1263

More details about data set preparation, model training, additional results and discussion, and supplementary data (PDF)

Polymer biodegradation data set (XLSX)

AUTHOR INFORMATION

Corresponding Author

Huichun Zhang − *Department of Civil and Environmental Engineering, Case Western Reserve University, Cleveland, Ohio 44106, United States;* orcid.org/0000-0002-5683-5117; Phone: (216) 368-0689; Email: hjz13@case.edu

Author

Chengrui Lin − *Department of Civil and Environmental Engineering, Case Western Reserve University, Cleveland, Ohio 44106, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.4c11282

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was mostly funded by the USDA Grant # 2020-67019-31019. The authors acknowledge helpful discussions with Drs. Kuan Huang, Susan Csiszar, Kathleen Mcdonough, and Ping Sun, as well as funding by Procter & Gamble toward the end of the project.

REFERENCES

(1) Plastics Europe. *Plastics − the Fast Facts 2023*, 2023. https://plasticseurope.org/knowledge-hub/plastics-the-fast-facts-2023/ (accessed 2024-09-03)

(2) Royal Society of Chemistry. Polymers in Liquid Formulations—Technical Report: A Landscape View of the Global PLFs Market; CPI, 2021. https://www.rsc.org/globalassets/22-new-perspectives/sustainability/liquid-polymers/rsc-polymer-liquid-formulations-technical-report.pdf (accessed 2024-08-01)

(3) Barron, A.; Sparks, T. D. Commercial Marine-Degradable Polymers for Flexible Packaging. *iScience* **2020**, *23* (8), 101353.

(4) Kyrikou, I.; Briassoulis, D. Biodegradation of Agricultural Plastic Films: A Critical Review. *J. Polym. Environ.* **2007**, *15* (2), 125−150.

(5) Alimi, O. S.; Farner Budarz, J.; Hernandez, L. M.; Tufenkji, N. Microplastics and Nanoplastics in Aquatic Environments: Aggregation, Deposition, and Enhanced Contaminant Transport. *Environ. Sci. Technol.* **2018**, *52* (4), 1704−1724.

(6) Ward, C. P.; Reddy, C. M. We Need Better Data about the Environmental Persistence of Plastic Goods. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117* (26), 14618−14621.

(7) Colwell, J.; Pratt, S.; Lant, P.; Laycock, B. Hazardous State Lifetimes of Biodegradable Plastics in Natural Environments. *Sci. Total Environ.* **2023**, *894*, 165025.

(8) Gross, R. A.; Kalra, B. Biodegradable Polymers for the Environment. *Science* **2002**, *297* (5582), 803−807.

(9) Aeschelmann, F.; Carus, M. Biobased Building Blocks and Polymers in the World: Capacities, Production, and Applications−Status Quo and Trends Towards 2020. *Ind. Biotechnol.* **2015**, *11* (3), 154−159.

(10) Lucas, N.; Bienaime, C.; Belloy, C.; Queneudec, M.; Silvestre, F.; Nava-Saucedo, J.-E. Polymer Biodegradation: Mechanisms and Estimation Techniques − A Review. *Chemosphere* **2008**, *73* (4), 429−442.

(11) Müller, R.-J.; Witt, U.; Rantze, E.; Deckwer, W.-D. Architecture of Biodegradable Copolyesters Containing Aromatic Constituents. *Polym. Degrad. Stab.* **1998**, *59* (1), 203−208.

(12) Selke, S.; Auras, R.; Nguyen, T. A.; Castro Aguirre, E.; Cheruvathur, R.; Liu, Y. Evaluation of Biodegradation-Promoting Additives for Plastics. *Environ. Sci. Technol.* **2015**, *49* (6), 3769−3777.

(13) Jayasekara, R.; Harding, I.; Bowater, I.; Lonergan, G. Biodegradability of a Selected Range of Polymers and Polymer Blends and Standard Methods for Assessment of Biodegradation. *J. Polym. Environ.* **2005**, *13* (3), 231−251.

(14) Haider, T. P.; Völker, C.; Kramm, J.; Landfester, K.; Wurm, F. R. Plastics of the Future? The Impact of Biodegradable Polymers on the Environment and on Society. *Angew. Chem., Int. Ed.* **2019**, *58* (1), 50−62.

(15) Eubeler, J. P.; Zok, S.; Bernhard, M.; Knepper, T. P. Environmental Biodegradation of Synthetic Polymers I. Test Methodologies and Procedures. *TrAC Trends Anal. Chem.* **2009**, *28* (9), 1057−1072.

(16) Albright, V. C.; Chai, Y. Knowledge Gaps in Polymer Biodegradation Research. *Environ. Sci. Technol.* **2021**, *55* (17), 11476−11488.

(17) Huang, K.; Prata, J. E.; Li, S.; Chen, Y.; Lin, C.; Baum, M.; Zhang, H. Development of a High-Throughput Fiber-Optic-Based Closed Bottle System for Evaluating the Ready Biodegradability of Water-Soluble Polymers. *ACS EST Water* **2024**, *4*, 2412.

(18) Strotmann, U.; Reuschenbach, P.; Schwarz, H.; Pagga, U. Development and Evaluation of an Online CO2 Evolution Test and a Multicomponent Biodegradation Test System. *Appl. Environ. Microbiol.* **2004**, *70* (8), 4621−4628.

(19) Ahn, B. D.; Kim, S. H.; Kim, Y. H.; Yang, J. S. Synthesis and Characterization of the Biodegradable Copolymers from Succinic Acid and Adipic Acid with 1,4-Butanediol. *J. Appl. Polym. Sci.* **2001**, *82* (11), 2808−2826.

(20) Takenouchi, S.; Takasu, A.; Inai, Y.; Hirabayashi, T. Effects of Geometric Structure in Unsaturated Aliphatic Polyesters on Their Biodegradability. *Polym. J.* **2001**, *33* (10), 746−753.

(21) Baba, T.; Tachibana, Y.; Suda, S.; Kasuya, K. Evaluation of Environmental Degradability Based on the Number of Methylene Units in Poly(Butylene n-Alkylenedionate). *Polym. Degrad. Stab.* **2017**, *138*, 18−26.

(22) Tachibana, Y.; Yamahata, M.; Ichihara, H.; Kasuya, K. Biodegradability of Polyesters Comprising a Bio-Based Monomer Derived from Furfural. *Polym. Degrad. Stab.* **2017**, *146*, 121−125.

(23) Wang, Y.; van der Maas, K.; Weinland, D. H.; Trijnes, D.; van Putten, R.-J.; Tietema, A.; Parsons, J. R.; de Rijke, E.; Gruter, G.-J. M. Relationship between Composition and Environmental Degradation of Poly(Isosorbide-Co-Diol Oxalate) (PISOX) Copolyesters. *Environ. Sci. Technol.* **2024**, *58* (5), 2293−2302.

(24) Nakayama, Y.; Watanabe, K.; Tanaka, R.; Shiono, T.; Kawasaki, N.; Yamano, N.; Nakayama, A. Synthesis, Properties, and Biodegradation of Sequential Poly(Ester Amide)s Containing γ-Aminobutyric Acid. *Int. J. Mol. Sci.* **2020**, *21* (10), 3674.

(25) Bading, M.; Olsson, O.; Kümmerer, K. Analysis of Environmental Biodegradability of Cellulose-Based Pharmaceutical Excipients in Aqueous Media. *Chemosphere* **2024**, *352*, 141298.

(26) Nakato, T.; Yoshitake, M.; Matsubara, K.; Tomida, M.; Kakuchi, T. Relationships between Structure and Properties of Poly(Aspartic Acid)s. *Macromolecules* **1998**, *31* (7), 2107−2113.

(27) Umare, S. S.; Chandure, A. S. Synthesis Characterization and Biodegradation Studies of Poly(Ester Urethane)s. *Chem. Eng. J.* **2008**, *142* (1), 65−77.

(28) Kim, J. R.; Thelusmond, J.-R.; Albright, V. C.; Chai, Y. Exploring Structure-Activity Relationships for Polymer Biodegradability by Microorganisms. *Sci. Total Environ.* **2023**, *890*, 164338.

(29) Wang, Y.; Van Putten, R.-J.; Tietema, A.; Parsons, J. R.; Gruter, G.-J. M. Polyester Biodegradability: Importance and Potential for Optimisation. *Green Chem.* **2024**, *26* (7), 3698−3716.

(30) Fransen, K. A.; Av-Ron, S. H. M.; Buchanan, T. R.; Walsh, D. J.; Rota, D. T.; Van Note, L.; Olsen, B. D. High-Throughput Experimentation for Discovery of Biodegradable Polyesters. *Proc. Natl. Acad. Sci. U.S.A.* **2023**, *120* (23), No. e2220021120.

(31) OECD. *Test No. 301: Ready Biodegradability*, 1992. https://www.oecd-ilibrary.org/environment/test-no-301-ready-biodegradability_9789264070349-en (accessed 2024-09-03).
(32) International Organization for Standardization. *Determination of the Ultimate Aerobic Biodegradability of Plastic Materials in an Aqueous Medium—Method by Measuring the Oxygen Demand in a Closed Respirometer*, 2019. https://www.iso.org/standard/70026.html (accessed 2024-09-03)
(33) Beran, E.; Hull, S.; Steininger, M. The Relationship Between the Chemical Structure of Poly(Alkylene Glycol)s and Their Aerobic Biodegradability in an Aqueous Environment. *J. Polym. Environ.* **2013**, *21* (1), 172−180.
(34) Ma, R.; Liu, Z.; Zhang, Q.; Liu, Z.; Luo, T. Evaluating Polymer Representations via Quantifying Structure−Property Relationships. *J. Chem. Inf. Model.* **2019**, *59* (7), 3110−3119.
(35) Kuenneth, C.; Rajan, A. C.; Tran, H.; Chen, L.; Kim, C.; Ramprasad, R. Polymer Informatics with Multi-Task Learning. *Patterns* **2021**, *2* (4), 100238.
(36) Yuan, W.; Hibi, Y.; Tamura, R.; Sumita, M.; Nakamura, Y.; Naito, M.; Tsuda, K. Revealing Factors Influencing Polymer Degradation with Rank-Based Machine Learning. *Patterns* **2023**, *4* (12), 100846.
(37) Huang, K.; Zhang, H. Classification and Regression Machine Learning Models for Predicting Aerobic Ready and Inherent Biodegradation of Organic Chemicals in Water. *Environ. Sci. Technol.* **2022**, *56*, 12755.
(38) Cheng, Y.; Zhang, K.; Huang, K.; Zhang, H. Meta-Analysis and Machine Learning Models for Anaerobic Biodegradation Rates of Organic Contaminants in Sediments and Sludge. *Environ. Sci. Technol.* **2024**, *58* (29), 12976−12988.
(39) Zhong, S.; Zhang, K.; Wang, D.; Zhang, H. Shedding Light on "Black Box" Machine Learning Models for Predicting the Reactivity of HO Radicals toward Organic Compounds. *Chem. Eng. J.* **2021**, *405*, 126627.
(40) Boethling, R. S.; Lynch, D. G. Biodegradation of US Premanufacture Notice Chemicals in OECD Tests. *Chemosphere* **2007**, *66* (4), 715−722.
(41) Zhong, S.; Guan, X. Count-Based Morgan Fingerprint: A More Efficient and Interpretable Molecular Representation in Developing Machine Learning-Based Predictive Regression Models for Water Contaminants' Activities and Properties. *Environ. Sci. Technol.* **2023**, *57* (46), 18193−18202.
(42) Jaeger, S.; Fulle, S.; Turk, S. Mol2vec: Unsupervised Machine Learning Approach with Chemical Intuition. *J. Chem. Inf. Model.* **2018**, *58* (1), 27−35.
(43) Xu, C.; Wang, Y.; Farimani, B. A Transformer-Based Language Model for Polymer Property Predictions. *npj Comput. Mater.* **2023**, *9* (1), 64.
(44) Calmon, A.; Silvestre, F.; Bellon-Maurel, V.; Roger, J.-M.; Feuilloley, P. Modelling Easily Biodegradability of Materials in Liquid Medium-Relationship Between Structure and Biodegradability. *J. Environ. Polym. Degrad.* **1999**, *7* (3), 135−144.
(45) Larrañaga, A.; Lizundia, E. A Review on the Thermomechanical Properties and Biodegradation Behaviour of Polyesters. *Eur. Polym. J.* **2019**, *121*, 109296.
(46) Tokiwa, Y.; Calabia, B.; Ugwu, C.; Aiba, S. Biodegradability of Plastics. *Int. J. Mol. Sci.* **2009**, *10* (9), 3722−3742.
(47) Chatterjee, M.; Banerjee, A.; De, P.; Gajewicz-Skretna, A.; Roy, K. A Novel Quantitative Read-across Tool Designed Purposefully to Fill the Existing Gaps in Nanosafety Data. *Environ. Sci.:Nano* **2022**, *9* (1), 189−203.
(48) Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A Next-Generation Hyperparameter Optimization Framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*; KDD '19; Association for Computing Machinery: New York, NY, USA, 2019; pp 2623−2631..
(49) Kuenneth, C.; Schertzer, W.; Ramprasad, R. Copolymer Informatics with Multitask Deep Neural Networks. *Macromolecules* **2021**, *54* (13), 5957−5961.
(50) Menzies, J.; Wilcox, A.; Casteel, K.; McDonough, K. Water Soluble Polymer Biodegradation Evaluation Using Standard and Experimental Methods. *Sci. Total Environ.* **2023**, *858*, 160006.
(51) McDonough, K.; Battagliarin, G.; Menzies, J.; Bozich, J.; Bergheim, M.; Hidding, B.; Kastner, C.; Koyuncu, B.; Kreutzer, G.; Leijs, H.; Parulekar, Y.; Raghuram, M.; Vallotton, N. Multi-Laboratory Evaluation of the Reproducibility of Polymer Biodegradation Assessments Applying Standardized and Modified Respirometry Methods. *Sci. Total Environ.* **2023**, *901*, 166339.
(52) Reuschenbach, P.; Pagga, U.; Strotmann, U. A Critical Comparison of Respirometric Biodegradation Tests Based on OECD 301 and Related Test Methods. *Water Res.* **2003**, *37* (7), 1571−1582.
(53) Kawai, F. Microbial Degradation of Polyethers. *Appl. Microbiol. Biotechnol.* **2002**, *58* (1), 30−38.
(54) Alford, D. D.; Wheeler, A. P.; Pettigrew, C. A. Biodegradation of Thermally Synthesized Polyaspartate. *J. Environ. Polym. Degrad.* **1994**, *2* (4), 225−236.
(55) Warren, R. A. J. MICROBIAL HYDROLYSIS OF POLYSACCHARIDES. *Annu. Rev. Microbiol.* **1996**, *50* (1), 183−212.
(56) Byrne, D.; Boeije, G.; Croft, I.; Hüttmann, G.; Luijkx, G.; Meier, F.; Parulekar, Y.; Stijntjes, G. Biodegradability of Polyvinyl Alcohol Based Film Used for Liquid Detergent Capsules: Biologische Abbaubarkeit Der Für Flüssigwaschmittelkapseln Verwendeten Folie Auf Polyvinylalkoholbasis. *Tenside, Surfactants, Deterg.* **2021**, *58* (2), 88−96.
(57) Nabeoka, R.; Suzuki, H.; Akasaka, Y.; Ando, N.; Yoshida, T. Evaluating the Ready Biodegradability of Biodegradable Plastics. *Environ. Toxicol. Chem.* **2021**, *40* (9), 2443−2449.
(58) Gaytán, I.; Burelo, M.; Loza-Tavera, H. Current Status on the Biodegradability of Acrylic Polymers: Microorganisms, Enzymes and Metabolic Pathways Involved. *Appl. Microbiol. Biotechnol.* **2021**, *105* (3), 991−1006.
(59) Zumstein, M. T.; Kohler, H.-P. E.; McNeill, K.; Sander, M. High-Throughput Analysis of Enzymatic Hydrolysis of Biodegradable Polyesters by Monitoring Cohydrolysis of a Polyester-Embedded Fluorogenic Probe. *Environ. Sci. Technol.* **2017**, *51* (8), 4358−4367.
(60) Mochizuki, M.; Hirami, M. Structural Effects on the Biodegradation of Aliphatic Polyesters. *Polym. Adv. Technol.* **1997**, *8* (4), 203−209.
(61) West, R. J.; Davis, J. W.; Pottenger, L. H.; Banton, M. I.; Graham, C. Biodegradability Relationships among Propylene Glycol Substances in the Organization for Economic Cooperation and Development Ready- and Seawater Biodegradability Tests. *Environ. Toxicol. Chem.* **2007**, *26* (5), 862−871.
(62) Beyler, C. L.; Hirschler, M. M. *Thermal Decomposition of Polymers*; Elsevier, 2002.
(63) Petinakis, E.; Liu, X.; Yu, L.; Way, C.; Sangwan, P.; Dean, K.; Bateman, S.; Edward, G. Biodegradation and Thermal Decomposition of Poly(Lactic Acid)-Based Materials Reinforced by Hydrophilic Fillers. *Polym. Degrad. Stab.* **2010**, *95* (9), 1704−1707.

https://doi.org/10.1021/acs.est.4c11282
Environ. Sci. Technol. 2025, 59, 1253−1263