

· 研究论文 ·

## 基于数据挖掘与领域知识分层特征构建的 Diels-Alder 聚氨酯性能预测与构效解析

袁李隽桀<sup>†</sup> 陈晓天<sup>†</sup> 游正伟<sup>\*</sup>

(东华大学材料科学与工程学院 先进纤维材料全国重点实验室 上海 201620)

**摘要** 含有 Diels-Alder (DA) 动态共价键的聚氨酯弹性体兼具优异的力学性能与热可逆自愈合能力, 是自修复弹性体领域的重要研究方向. 本工作基于文献数据驱动方法, 围绕 DA 聚氨酯弹性体(DAPU)拉伸强度、断裂伸长率和愈合效率这 3 项性能的构效关系开展研究. 通过对相关文献中配方与性能数据的系统筛选, 构建了涵盖组成、化学特性和结构机理 3 个层次的参数体系. 进一步选取支持向量回归(SVR)、极限梯度提升(XGBoost)和高斯过程回归(GPR)分别建立了 3 项性能的定量解析模型, 其决定系数( $R^2$ )分别达到 0.76、0.66 和 0.81. 结果表明, 所构建的特征体系能够较好表征 DAPU 弹性体主要结构变量与性能之间的关系. 夏普利加性解释(SHapley Additive exPlanations, SHAP)分析进一步表明, 软段结晶性和软段分子量是影响力学强度与愈合效率的最关键参数, DA 功能单体含量是调控延展性的核心变量, 3 项性能之间呈现出复杂的非单调相关关系. 上述结果可为深入理解 DAPU 弹性体的构效关系, 指导配方设计提供参考.

**关键词** 聚氨酯; Diels-Alder 动态键; 自愈合弹性体; 机器学习; 构效关系; 数据驱动

**引用:** 袁李隽桀, 陈晓天, 游正伟. 基于数据挖掘与领域知识分层特征构建的 Diels-Alder 聚氨酯性能预测与构效解析. 高分子学报, doi: 10.11777/j.issn1000-3304.2026.26122.

**Citation:** Yuanli, J. J.; Chen, X. T.; You, Z. W. A literature data-driven machine learning study of structure-property relationship in Diels-Alder polyurethane elastomers based on hierarchical domain-knowledge features. *Acta Polymerica Sinica* (in Chinese), doi: 10.11777/j.issn1000-3304.2026.26122.

含有 Diels-Alder (DA) 动态共价键的聚氨酯弹性体(DAPU)是一类在聚氨酯网络中引入呋喃-马来酰亚胺 DA 动态共价键的热可逆共价自适应网络(CAN)材料, 兼具聚氨酯弹性体良好的力学性能与无需外援修补剂的本征自愈合能力, 在柔性电子、软体机器人和可穿戴器件等领域具有重要应用前景<sup>[1-8]</sup>. DAPU 通常由软段多元醇、硬段组分(异氰酸酯与扩链剂)以及 DA 功能单元构成, 聚集形成微相分离的软硬两相结构. 其中硬相主要提供物理

交联和力学支撑, 软相则主导链段运动性和材料延展性. DA 单元在升温至约 100~150 °C 时会发生逆 Diels-Alder 反应解离, 降温后又重新成键, 这一特性赋予了材料良好的热可逆愈合与重复加工能力<sup>[9-11]</sup>. 因此近年来, 基于 DA 动态键的聚氨酯自修复材料持续受到关注. 以“polyurethane elastomer”、“dynamic polyurethane”和“Diels-Alder polyurethane”为关键词在 Web of Science 中进行检索可见, 相关研究数量总体呈增长趋势, 如图 1 所示.

2026-04-25 收稿, 2026-05-27 录用, 网络出版.

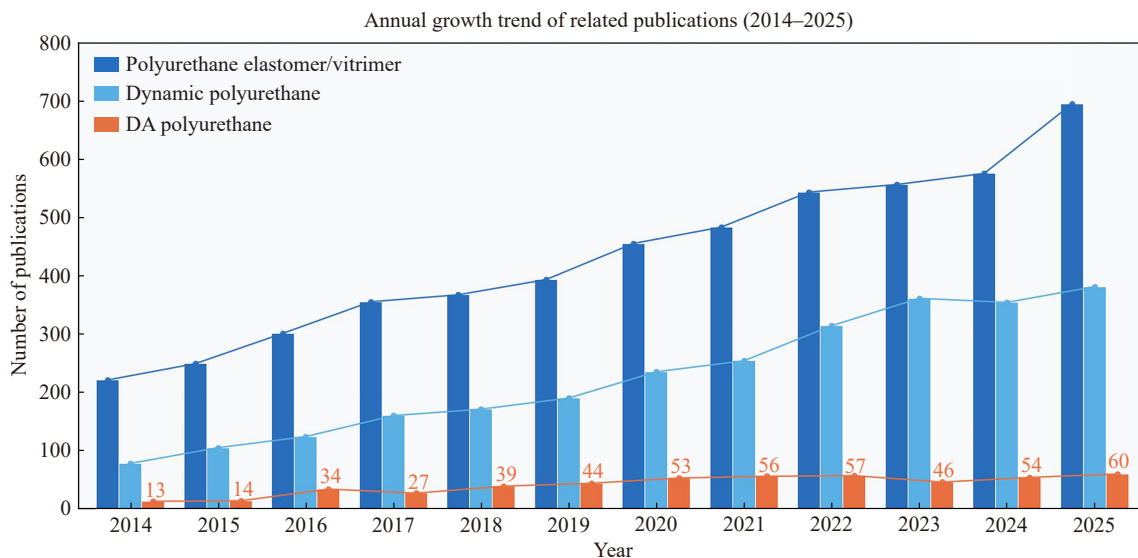
基金项目: 国家自然科学基金(基金号 U25A20251, 52473004, 52173117), 长三角科技创新共同体联合攻关项目(项目号 Basic Research Program of Jiangsu, 2025CSJZN00800, BK20254004).

<sup>†</sup> 共同第一作者.

<sup>\*</sup> 通信联系人, E-mail: zyou@dhu.edu.cn

本文附有电子支持信息, 与正文一并刊登在本刊网站(www.gfzxb.org).

doi: 10.11777/j.issn1000-3304.2026.26122; CSTR: 32057.14.GFZXB.2026.7632



**Fig. 1** Annual growth trend of related publications for polyurethane elastomers, dynamic polyurethane and DA polyurethane (2014–2025).

相较于其他自修复聚氨酯体系，DAPU弹性体因无需催化剂、愈合温度可调及力学可设计空间大，已成为该领域的重要研究方向。DAPU弹性体的性能优化通常面临力学强度与自愈合效率之间的权衡：较高的拉伸强度往往依赖于更充分的微相分离和更高的交联密度，而较高的愈合效率则要求链段在愈合温度下具有充分的运动性，二者对网络拓扑结构的要求相互制约<sup>[12–16]</sup>。此外，多元醇分子量与类型、异氰酸酯种类、DA单体含量、NCO/OH比值等参数共同构成了复杂的高维配方空间，单纯依赖经验判断或实验试错寻求最优配方成本高昂，因此有必要建立面向DAPU弹性体的定量构效关系模型，为配方设计提供理论依据。

机器学习已逐渐成为高分子材料构效关系研究的重要工具，已有报道展示在聚氨酯领域引入机器学习具有重要意义<sup>[17–26]</sup>。包括聚合物多变量关系建模、硬质聚氨酯泡沫性能预测、水性聚氨酯建模以及聚氨酯力学性能预测等方面均有展现出基于大数据方法解析的潜力。而针对动态聚氨酯这一复杂的共价网络体系的研究目前仍较为缺乏，这是本工作的重要切入点。然而，DAPU弹性体相关性能数据通常分散在不同文献报道中，不同研究在应变速率、试样尺寸和测试条件等方面存在显著差异，使文献数据具有较强异质性，增加了定量建模的难度。此外，仅采用化学计量组成等基础描述符，往往难以充分表征DAPU体

系中由链段结构、微相分离和动态键局部环境共同决定的复杂构效关系，从而限制了文献数据的有效利用<sup>[27,28]</sup>。目前，基于异质文献数据对DAPU弹性体力学性能与自愈合性能开展可解释的机器学习解析的研究仍较为有限。

基于此，本工作围绕“DAPU弹性体的力学性能与自愈合性能能否通过异质文献数据建立定量解析模型”的问题展开，基于系统的异质文献挖掘，构建由领域知识引导的3层特征框架，分别采用支持向量回归、XGBoost和高斯过程回归建立拉伸强度、断裂伸长率和自愈合效率的定量解析模型。在此基础上，结合SHAP分析揭示关键特征对不同性能指标的作用方向及相对重要性。

## 1 DA聚氨酯弹性体

### 1.1 DAPU弹性体的结构特征与主要分类变量

DAPU弹性体的结构多样性主要体现在软段类型、异氰酸酯类型和DA连接单元的局部化学环境等方面，这些因素共同决定了材料的微相结构及其力学性能与自愈合行为。根据DA功能单元在网络中的拓扑位置，DAPU可分为主链型和侧链型两类，前者中DA加合物构成主链骨架，后者中DA加合物位于侧链。根据DA连接处局部化学环境，DA连接体又可进一步分为氨基键型和脲键型，分别对应以糠醇和糠胺为扩链剂的体系，两者的氢键作用强度及对软段链运动性的影

响方面存在明显差异<sup>[29]</sup>。此外，软段多元醇通常分为聚酯型和聚醚型，前者主要包括聚己内酯(PCL)、聚己二酸丁二醇酯(PBA)等，后者主要包括聚四亚甲基醚二醇(PTMG)、聚丙二醇(PPG)等，2类软段在微相分离倾向、结晶能力和链段柔性等方面存在显著差异<sup>[30]</sup>。现有文献报道表明，DAPU弹性体中常用的软段以PTMG和PPG为主，聚酯型软段如PCL、PBA等亦有报道，不同软段类型体系的力学性能与愈合性能差异显著<sup>[31,32]</sup>。因此，软段类型及其相关结构特征是后续构建DAPU弹性体构效关系模型时需要重点考虑的关键变量。

## 1.2 DAPU弹性体力学性能与自愈合效率的结构基础

DAPU弹性体中主要涉及2个维度：交联密度和链段运动性。前者主要受硬段含量和DA单体用量影响，后者则主要取决于软段分子量及其化学类型。交联密度升高时，硬段微区和物理交联点增多，有利于提高材料的拉伸强度；但同时增大了拓扑约束，抑制愈合温度下链段的跨界面扩散，从而不利于自愈合。相反，链段运动性提高时，软段链扩散加快，愈合效率提升，但有效交联密度降低，材料拉伸强度随之下降。

在上述框架中，软段结晶性具有独特作用。易结晶软段可在室温下形成结晶微区，提供额外的物理交联，从而增强材料的力学性能，如PCL、PTMG；而在修复温度下，这些结晶结构发生熔融，链段活动性提高，有助于界面链扩散与动态键的重新成键。因此，这一室温结晶增强模量、升温熔融促进扩散的时间尺度分工机制，使得可结晶软段体系在力学性能与自愈合效率的权衡中具有独特优势。

总体来看，DAPU弹性体的配方参数与力学性能、自愈合效率之间并不存在简单的线性对应关系，而是受到交联密度、链段运动性、微相分离及软段结晶等多种结构因素耦合作用的共同控制。因此，有必要借助数据驱动方法对其构效关系进行进一步解析。

## 2 DAPU性能的数据驱动分析

### 2.1 数据集构建

以“Diels-Alder polyurethane”、“furan maleimide polyurethane”和“thermoreversible self-

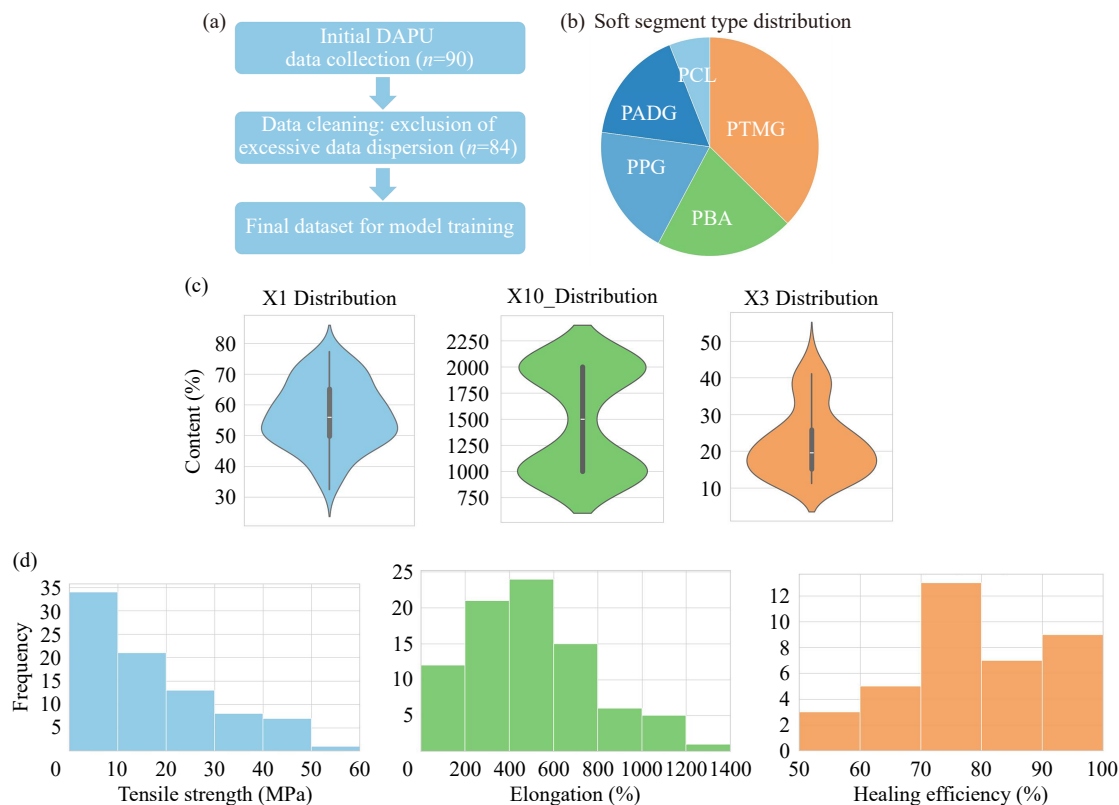
healing polyurethane elastomer”为检索词，在Web of Science、Scopus和Google Scholar上进行系统文献检索(截止时间为2025年6月)，并结合关键综述的前向与后向引文追踪补充文献来源。数据的纳入标准为：以呋喃-马来酰亚胺DA加合物为动态交联单元，同时报告多元醇类型与分子量、异氰酸酯类型、DA单体种类及NCO/OH摩尔比，基于初步筛选得到的文献，至少报告拉伸强度、断裂伸长率或自愈合效率中的一项定量数据。手动提取约满足最低报告要求的DAPU配方与性能数据。纳入文献需同时报告多元醇类型与分子量、异氰酸酯类型、DA单体种类、NCO/OH摩尔比以及定量性能指标。

原始数据集共包含92个配方样本。经数据清洗和筛选后，构建了2个建模数据集。如图2(a)所示，清洗过程中共剔除8条记录(包括重复数据、关键特征缺失及测试条件不明确的样本)，得到力学建模集合A( $N=84$ )。在此基础上进一步筛选出有效报告自愈合效率的样本：(1)愈合效率以定量形式报告；(2)明确给出愈合温度和愈合时间。最终得到愈合效率建模集合B( $n=37$ )，集合B为集合A的子集。

如图2(b)所示，集合A中软段组成以PTMG为主约占37%，聚己二酸丁二醇酯(PBA)和聚己二酸乙二醇酯(PADG)各约占19%，PCL约占6%，构成相对稀疏的子群体。如图2(c)所示，主要连续描述符具有较宽的覆盖范围。其中软段质量分数为0.32~0.77(X1)，硬段含量也在较大的范围空间内波动(X3)，说明训练数据集能有效覆盖DAPU配方空间中化学多样性丰富的区域。本文后续对特征进行了从X1至X16的对应编号，每个特征均有独立编号(见表1)。

图2(d)进一步表明，本文收集的DAPU弹性体数据中，拉伸强度分布在0.63~50.3 MPa、断裂伸长率在96%~1364%、自愈合效率在53.0%~99.0%。上述结果说明，当前数据集在配方和性能2个维度上均具有较宽的覆盖范围，有利于后续模型训练及泛化能力评价。

为解决数据异质性影响，所有连续输入特征和输出目标在建模前均经过Z-score标准化，以消除不同特征量纲差异对模型的影响。同时为避免数据泄露，标准化过程嵌入了交叉验证流程。在模型选用方面选择了自带预测不确定性的高斯



**Fig. 2** Statistical characteristics of the literature-derived DAPU dataset. (a) Data curation workflow from initial collection to the final training set ( $n=84$ ); (b) Distribution of soft-segment polyol types; (c) Violin plots of key descriptors: soft-segment content (X1), hard-segment content (X3), and  $Soft_{M_w}$  (X10); (d) Histograms of tensile strength, elongation at break, and self-healing efficiency.

过程回归(GPR)及对输入尺度不敏感的树集成模型, 并配合留一法交叉验证评估泛化性能。

## 2.2 特征参数的构建

为克服纯化学计量描述符的信息局限性, 本工作在组成-结构-性质-性能关系(CSPP)框架下<sup>[33-36]</sup>, 构建了一个基于物理信息丰富程度逐级递增的3个层特征参数体系, 用于编码DAPU的配方, 同时对于自愈体系, 还融入了2个过程特征. 具体的特征设计见表1 (具体特征描述见电子支持信息).

为弥补上述不足, 第二层进一步引入4个反映组分化学类别的分类变量(Identity). 这4个变量能够编码常规计量参数难以体现的化学差异与热力学特征. 例如: 脲键型DA连接单元由于N-H供体密度更高, 其氢键作用力通常强于氨基型, 从而可能对性能产生不同影响。

在此基础上, 第三层进一步引入与分子量和链结构密切相关的5个机理特征(Mechanism). 该类特征将高分子物理中的结构信息转化为可直接输入模型的标量参数, 从而弥补剂量特征和化学

类别特征在物理信息表征上的不足。

对于自愈合效率模型, 考虑到不同文献来源的愈合方案差异较大, 在上述14个配方特征基础上, 进一步引入愈合温度和愈合时间两个愈合参数, 以降低不同修复条件差异对愈合效率目标的干扰。

基于上述特征体系, 进一步计算了各特征参数与3项性能指标间的Pearson相关系数矩阵. 如图3所示, 选取6个主要连续配方特征与3项性能指标, 构建两两关系矩阵, 并结合散点分步矩阵和分组核密度分布图考察其潜在线性、非线性关系和分群特征。

Pearson相关性分析表明, 各特征与目标性能的最大绝对线性相关分别为: 拉伸强度 $|r|_{\max} = 0.653$  (Hard\_Content), 断裂伸长率 $|r|_{\max} = 0.425$  (DA\_Content), 自愈合效率 $|r|_{\max} = 0.777$  ( $Soft_{M_w}$ ). 尽管部分特征与目标性能之间已表现出一定程度的线性关联, 但DAPU体系的构效关系仍较为复杂. Hard\_Content与拉伸强度的趋势随Soft\_Cryst定义的子群体不同而变化; DA\_Content与断裂

**Table 1** Three-tier feature parameter system for DAPU elastomers.

ID	Feature	Level	Formula	Physical basis
X1	Soft_Content	Dosage	$X1 = m_{\text{polyol}}/m_{\text{total}}$	Soft-segment mass fraction; governs chain mobility and microphase separation degree <sup>[35]</sup>
X2	DA_Content	Dosage	$X2 = m_{\text{DA monomer}}/m_{\text{total}}$	DA group mass fraction; determines density of thermoreversible crosslinking sites
X3	Hard_Content	Dosage	$X3 = (m_{\text{iso}} + m_{\text{extender}})/m_{\text{total}}$	Hard-segment mass fraction; controls strength of physical crosslinking domains
X4	R_Ratio	Dosage	$X4 = \Sigma n_{\text{NCO}}/\Sigma n_{\text{active H}}$	NCO/OH molar ratio; governs network integrity and crosslinking density
X5	Add_Crosslink	Dosage	$X5 = m_{\text{crosslinker}}, f > 2/m_{\text{total}}$	Permanent crosslinker content; constructs dual-network architecture
X6	Network_Topology	Identity	0 = backbone DA; 1 = side-chain DA	Topological position of DA bond; determines fracture mode and healing kinetics <sup>[29]</sup>
X7	DA_Linker_Type	Identity	0 = no DA linker; 1 = urea-type; 2 = carbamate-type (furfuryl alcohol)	Local hydrogen-bonding environment at DA junction; affects chain dynamics and activation energy
X8	Polyol_Class	Identity	0 = polyether; 1 = polyester	Soft-segment polarity; influences phase separation driving force and chain diffusion rate <sup>[25]</sup>
X9	Iso_Class	Identity	0 = aliphatic; 1 = aromatic	Isocyanate type; determines hard-segment $\pi$ - $\pi$ stacking and hydrogen-bonding network strength
X10	Soft_Mw	Mechanism	$\text{Soft } M_w = \Sigma(m_{\text{polyol},i} \cdot M_i)/\Sigma m_{\text{polyol},i}$	Average soft-segment molecular weight; controls crosslink spacing $M_c$ and chain reptation rate <sup>[15]</sup>
X11	Constraint_Factor	Mechanism	$X3/\text{Soft } M_w$	Topological constraint index; quantifies hard-segment restriction per unit soft-segment length
X12	Hard_Symmetry	Mechanism	1 = Symmetric; 0 = asymmetric	Hard-segment molecular symmetry; affects hard-domain ordering and crystallization propensity <sup>[37]</sup>
X13	Soft_Cryst	Mechanism	1 = Crystallizable soft segment and $M_w \geq 600$ ; 0 = otherwise (g/mol)	Soft-segment crystallizability; provides physical crosslinking points and thermally driven healing switch <i>via</i> crystal melting-recrystallization <sup>[37,38]</sup>
X14	Synergy_Feature	Mechanism	$X1 \times \text{Soft\_Cryst} \times \text{Hard\_Symmetry}$	Structural synergy factor; identifies high-performance systems with both ordered soft and hard segments
X15	Healing_Temp	Process	$T_{\text{healing}} (\text{°C})$	Thermal healing temperature; governs retro-DA dissociation and chain diffusion kinetics during healing <sup>[9-11]</sup>
X16	Healing_Time	Process	$t_{\text{healing}} (\text{h})$	Thermal healing duration; determines sufficient time for interface chain interdiffusion and DA re-bonding

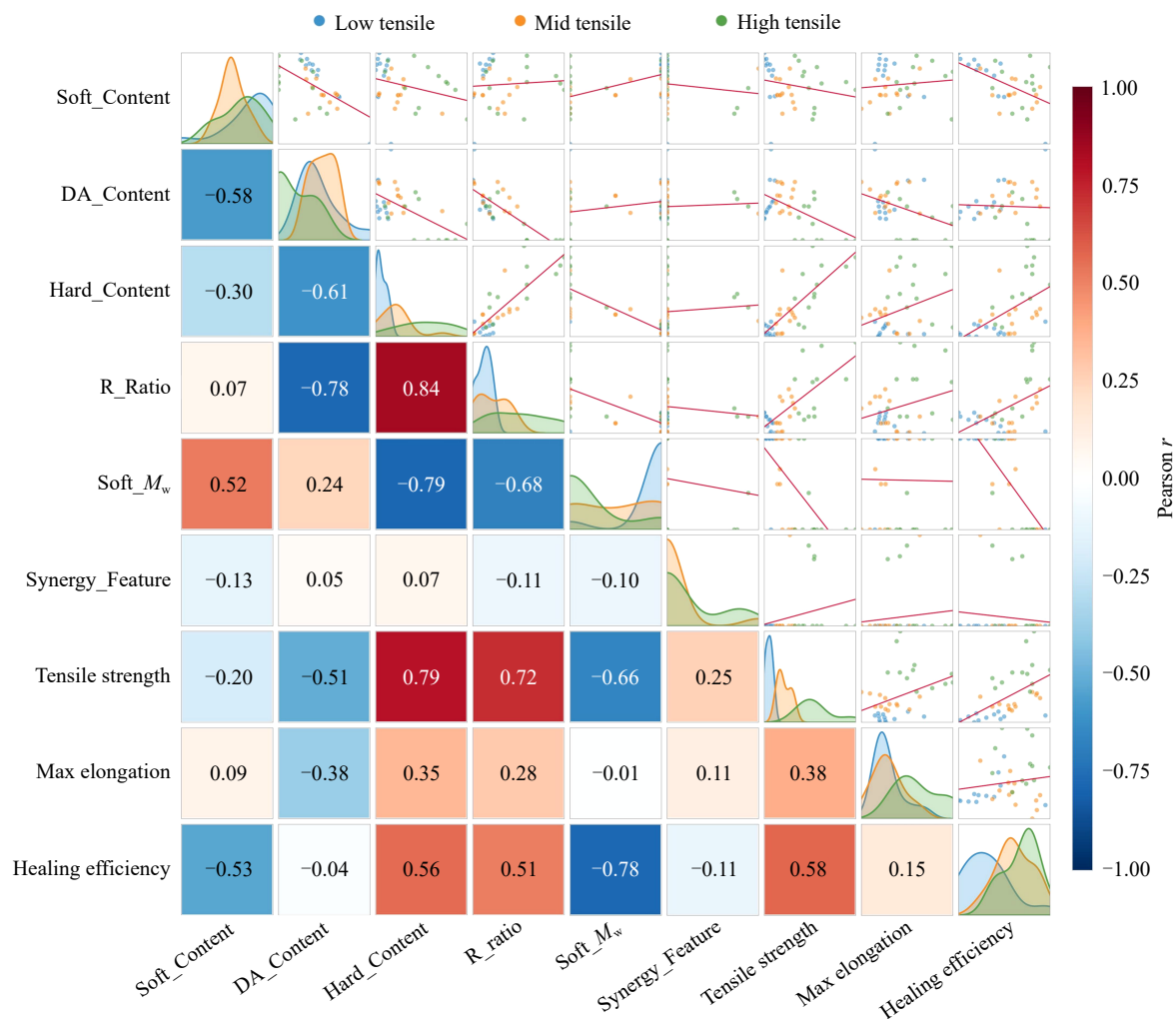
伸长率之间则呈现明显的非单调关系，即在适中含量范围内对应较高延展性，而当DA含量继续升高后断裂伸长率显著下降。上述结果表明，简单线性描述难以充分刻画DAPU弹性体的构效关系，这也为后续引入机理层特征和非线性模型提供了依据。

### 2.3 机器学习解析模型的构建与评价

为建立DAPU弹性体3项性能的定量预测模型，系统比较了8种候选的机器学习算法，包括支持向量回归(SVR)、高斯过程回归(GPR)、随

机森林(RF)、套索回归(LASSO)、岭回归(Ridge)、极端梯度提升树(XGBoost)、多层感知机(MLP)和多元线性回归(MLR)等<sup>[39-43]</sup>。以留一法交叉验证(LOOCV)的决定系数 $R^2$ 为主要评价指标，分别筛选出适用于不同性能指标的最优算法。最终，拉伸强度采用带RBF核的SVR模型，断裂伸长率采用XGBoost模型，自愈合效率采用带Matérn 3/2核的GPR模型(具体参数见电子支持信息图SI)。

考虑到数据集规模有限(集合A,  $N=84$ , 集合



**Fig. 3** Pairwise relationship matrix of continuous formulation features and performance metrics of DAPU elastomers. Diagonal: class-stratified KDEs by tensile strength (low/mid/high). Lower triangle: Pearson  $r$  values. Upper triangle: scatter plots with linear regression guides.

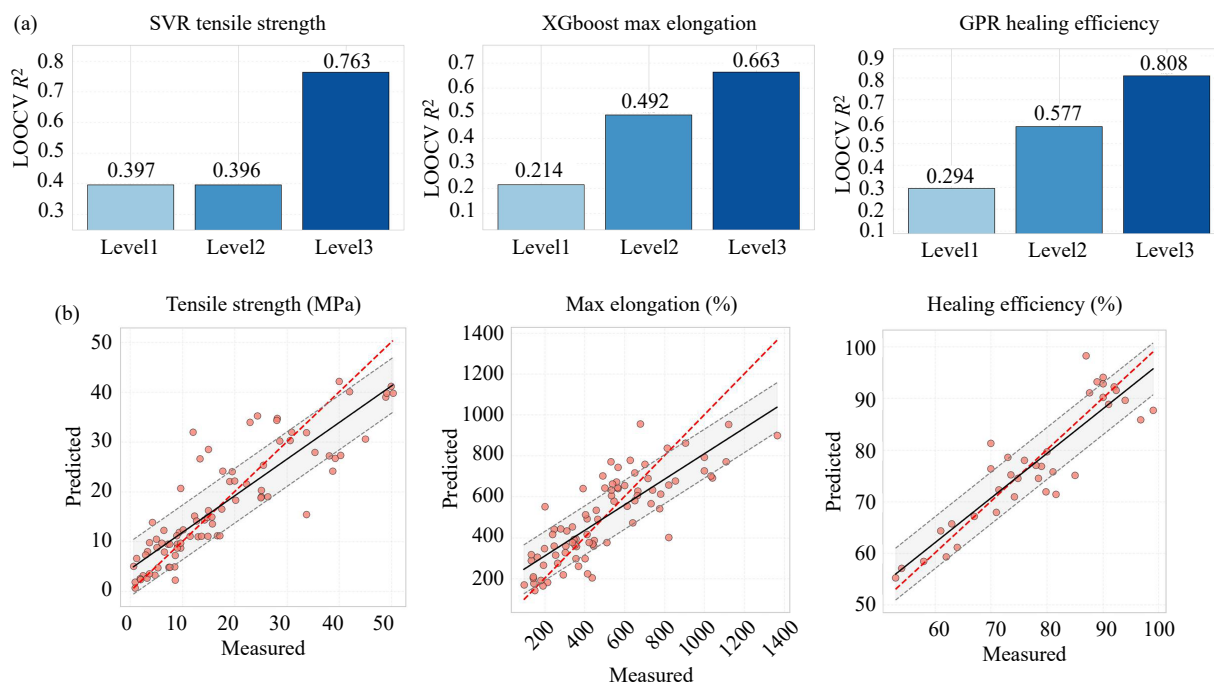
B,  $N=37$ ), 本工作采用留一法交叉验证 LOOCV 方案对模型性能进行评估, 并以 LOOCV  $R^2$  作为模型泛化性能的评价指标<sup>[44]</sup>. 所有计算在 Python 3.9.6 环境中完成 (scikit-learn 1.6.1, XGBoost 2.1.4), 模型可解释性分析采用 SHAP, 树模型使用 TreeExplainer, 核方法使用 KernelExplainer<sup>[45-48]</sup>.

不同特征层级下模型性能的变化如图 4(a) 所示. 仅使用第一层剂量特征时, 拉伸强度、断裂伸长率和自愈合效率的 LOOCV  $R^2$  分别为 0.40、0.21 和 0.29, 说明仅依赖化学计量组成难以从异质文献数据中提取可靠的构效信息. 引入第二层化学特性特征后, 断裂伸长率与自愈合效率  $R^2$  分别提升至 0.49 和 0.58, 而拉伸强度  $R^2$  仍维持在 0.40, 表明化学类别信息可有效解析软段运动性相关的性能差异, 但仍不足以刻画强度所依赖的

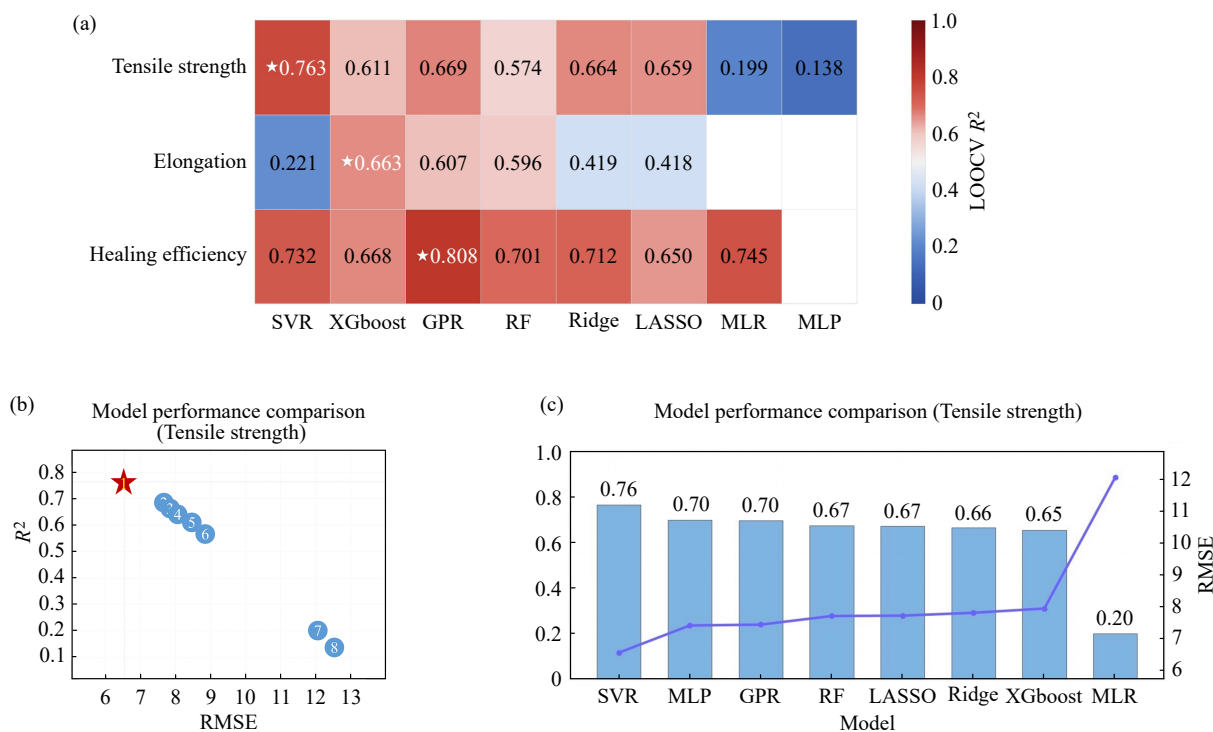
的分子量和链结构效应. 进一步引入第三层结构机理特征后, 3 项性能  $R^2$  分别达到 0.76、0.66 和 0.81, 较之前有明显提升, 表明将高分子物理相关信息转化为可计算特征, 对提升 DAPU 弹性体构效关系模型的预测能力具有重要作用.

需要说明的是, 3 项任务的  $R^2$  数值不进行完全等价的横向比较. 从变量的分布来看, 自愈合率的分布符合正态分布, 而拉伸强度和断裂伸长率呈正偏态, 数据的偏态分布增大了回归模型在全值域的拟合难度, 是自愈合效率  $R^2$  高于其他 2 个目标的直接统计原因. 并且因为引入了自愈合的 2 项工艺信息, 一定程度上弥补了跨文献数据来源带来的噪声.

除  $R^2$  外, 进一步以拉伸强度预测模型为例, 对多种候选模型的交叉验证均方根误差 (RMSE)



**Fig. 4** Model performance evaluation and descriptor-tier dependence of the machine learning models for DAPU elastomers: (a) LOOCV  $R^2$  values for the three target properties as a function of descriptor hierarchy. Bar shades represent the hierarchical inclusion of Tier-1 (dosage), Tier-2 (chemical identity), and Tier-3 (mechanistic) features. (b) Parity plots comparing LOOCV-predicted versus measured values. Red dashed lines denote the  $y = x$  ideal; black solid lines and shaded regions indicate linear fits and confidence intervals, respectively.



**Fig. 5** Comprehensive evaluation of machine learning model performance for predicting self-healing polyurethane. (a) Heatmap of LOOCV  $R^2$  scores across eight candidate models and three target properties with starred cells marking the best-performing model for each property; (b) Scatter plot of  $R^2$  versus RMSE for tensile strength prediction, where the red star denotes the optimal model (SVR); (c) Bar-and-line chart for tensile strength, where blue bars show LOOCV  $R^2$  values ranked in descending order and the purple line traces the corresponding RMSE trend across models.

进行了比较,如图5所示.结果表明,SVR模型不仅具有最高的 $R^2$ ,同时在RMSE指标上也表现出更低的整体误差水平,说明其在拟合精度与泛化稳定性之间取得了较好的平衡.相比之下,线性模型因难以表征DAPU体系中的非线性构效关系而误差较大,树模型虽能刻画局部阈值效应,但在当前样本规模下,连续平滑变化趋势的捕捉仍略逊于SVR.

3种不同的算法分别在3个预测任务中取得最优结果,这与各性能所对应的构效关系特征具有一定一致性:拉伸强度变化趋势相对平滑,适合SVR的全局连续非线性回归;断裂伸长率呈Soft\_Cryst转变和DA含量临界值附近的阈值行为,适合XGBoost的分段非线性关系的刻画;愈合效率样本量较小、且不同文献间愈合方案差异较大,因此GPR在处理小样本问题及表征预测不确定性方面更具优势.

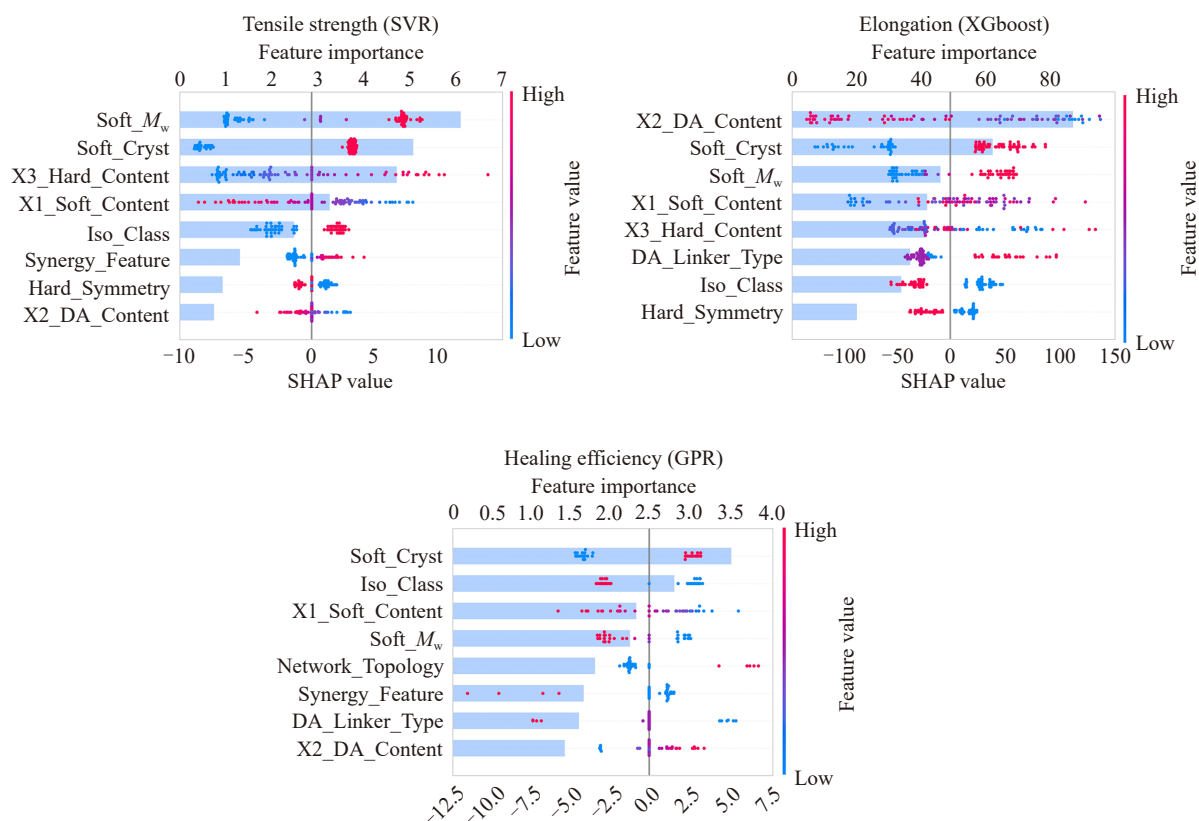
上述结果表明,DAPU弹性体不同性能指标所对应的构效关系形式存在明显差异,因此需要采用与其数据特征相匹配的建模策略.需要进一

步指出的是,断裂伸长率模型的LOOCV  $R^2$ 为0.66,虽略低于拉伸强度和自愈合效率,但在当前异质文献数据背景下仍具有合理性.断裂伸长率对试样缺陷、应变速率、测试温度及哑铃样条尺寸等实验条件更为敏感,不同研究之间测试协议差异会被直接放大到最终响应中.同时,断裂伸长率受链段取向、微相分离程度以及局部缺陷演化等多因素耦合作用,其构效关系本身也更容易表现出阈值型和非单调行为.因此,在样本量有限且跨文献条件难以完全统一的前提下,现有 $R^2$ 已说明模型能够提取主导延展性变化的关键信息.

## 2.4 SHAP特征重要性分析

各目标性能的SHAP分析结果如图6所示.整体上,3项目标性能分别受不同特征主导,且部分关键变量对不同性能表现出相互制约的作用趋势.各特征贡献方向与高分子物理预期相符<sup>[49,50]</sup>.

由图可知,Soft\_Cryst(软段结晶性)在3项性能中均呈正向贡献<sup>[51]</sup>,其作用前提在于PCL、



**Fig. 6** Model interpretability *via* SHAP analysis. SHAP beeswarm plots illustrating the local influence of the 8 most significant descriptors for tensile strength, elongation, and healing efficiency. Point colors represent feature magnitudes (blue: low; red: high), with horizontal displacement indicating the contribution to the respective model prediction.

PTMG 等软段属于低熔点可逆结晶，可在愈合温度下完全熔融，有别于高熔点永久结晶对链段运动的持续抑制。力学层面，结晶微区作为可逆物理交联点提升承载能力；拉伸过程中发生应力诱导熔融，将机械能转化为相变潜热，延缓断裂并维持高伸长率。自愈合层面，加热时晶区熔融，为 DA 键重建提供初始接触条件；冷却时软段再结晶在愈合界面原位形成新的物理交联点，将重建的共价网络锁定，直接提升愈合强度保留率。此外，软段结晶性越高，软硬段微相分离越充分，DA 键断裂后官能团更容易相互靠近并完成重建。因此，软段结晶性通过热激活、再结晶锁定与微相结构优化的协同作用，成为 DAPU 体系中兼顾力学与自愈合性能的核心结构变量。

Soft\_ $M_w$  (软段平均分子量)对 3 项性能表现出典型的权衡特征。对拉伸强度呈负向贡献：Soft\_ $M_w$  增大，则交联点间距增大，以至有效交联密度降低，与经典橡胶弹性理论一致；但对断裂伸长率和自愈合效率均呈正向贡献：较长软段链提高链段柔顺性与变形能力，加速界面互扩散与重新缠结<sup>[52]</sup>。综合来看，Soft\_ $M_w$  的增大倾向于“软化”材料，即不利于强度提升，但有助于延展与愈合能力的改善，因此实际配方设计中需根据目标性能侧重点进行调节。

DA\_Content (DA 单体含量)对断裂伸长率呈主导负向贡献，表明 DA 含量升高会增强网络约束并降低链段延展性。与此同时，DA\_Content 对拉伸强度通常呈正向贡献(更强的网络约束有助于强度提升)，进一步印证了强度-延展性之间的权衡。

Hard\_Content (硬段含量)与 Iso\_Class (异氰酸酯类型)主要影响拉伸强度。硬段含量增加有助于提高硬相微区比例，增强材料承载能力；芳香族异氰酸酯(相对于脂肪族)凭借更优越的  $\pi$ - $\pi$  堆积和氢键网络，有利于硬段微区有序化，从而提升材料强度。然而，Iso\_Class 对自愈合效率呈负向贡献。这表明，芳香族硬段在愈合温度下对链段运动的约束较强，不利于界面扩散与 DA 键动态交换的进行<sup>[53,54]</sup>。上述结果表明，硬段结构在提升力学强度的同时，也可能对愈合行为形成制约，体现出强度与愈合效率之间在结构层面的内在矛盾。

需要指出的是，当前数据集虽然已能够支持建立具有一定解释能力的定量模型，但样本覆盖仍然有限，对于部分机理的解释和影响关系依然不够清晰。后续研究可通过扩大样本规模、提高稀疏子群体的覆盖度，并引入连续结晶倾向指数等更精细的机理特征，进一步提升模型的表征能力与泛化性能<sup>[55]</sup>。

### 3 结论与展望

基于异质文献数据，围绕 DA 聚氨酯弹性体的拉伸强度、断裂伸长率和自愈合效率 3 项关键性能，构建了包含剂量特征、化学类别特征和机理特征在内的 3 层参数体系，并分别建立了对应的机器学习解析模型。结果表明，3 类性能的 LOOCV  $R^2$  分别达到 0.76、0.66 和 0.81，说明在文献数据存在跨实验室差异、样本规模有限的条件下，所构建的特征体系仍能够较好捕捉 DAPU 弹性体主要结构变量与性能之间的关联规律。

进一步的 SHAP 分析表明，Soft\_Cryst、Soft\_ $M_w$  和 DA\_Content 是决定 DAPU 性能权衡关系的关键变量。其中，Soft\_Cryst 是少数能够同时对强度、延展性和自愈合效率产生协同正向作用的特征；Soft\_ $M_w$  增大通常有利于链段运动、延展与愈合，但会削弱有效交联约束并降低强度；DA\_Content 提高则有助于增强网络约束，却可能抑制材料延展性。3 者之间存在内在协同关系，配方优化应以 3 变量协同正贡献区间为目标，围绕“可逆网络密度-链段运动性-微相有序结构”之间的平衡开展多目标协同优化，而非单一追求某一组分的最大化。

需要指出的是，当前研究仍受到样本规模、实验条件异质性以及部分机理特征离散化表达的限制。未来可从以下几个方面进一步拓展：持续扩大文献样本覆盖范围；引入连续化机理描述符，如结晶度、氢键强度指标和相分离表征参数，以提升模型对真实结构差异的刻画能力；推动自愈合测试条件与力学测试协议的标准化，从源头降低文献数据异质性。随着数据积累和特征表达能力的提升，DAPU 弹性体的机器学习模型有望从“规律解释工具”进一步发展为“配方反向设计工具”，为高性能自愈合聚氨酯材料的理性设计提供更可靠支持。

## REFERENCES

- 1 Kloxin, C. J.; Scott, T. F.; Adzima, B. J.; Bowman, C. N. Covalent adaptable networks (CANs): a unique paradigm in cross-linked polymers. *Macromolecules*, **2010**, 43(6), 2643–2653.
- 2 Rus, D.; Tolley, M. T. Design, fabrication and control of soft robots. *Nature*, **2015**, 521(7553), 467–475.
- 3 Rogers, J. A.; Someya, T.; Huang, Y. G. Materials and mechanics for stretchable electronics. *Science*, **2010**, 327(5973), 1603–1607.
- 4 Chortos, A.; Liu, J.; Bao, Z. N. Pursuing prosthetic electronic skin. *Nat. Mater.*, **2016**, 15(9), 937–950.
- 5 Wang, S. Y.; Urban, M. W. Self-healing polymers. *Nat. Rev. Mater.*, **2020**, 5(8), 562–583.
- 6 Briou, B.; Améduri, B.; Boutevin, B. Trends in the Diels-Alder reaction in polymer chemistry. *Chem. Soc. Rev.*, **2021**, 50(19), 11055–11097.
- 7 王瑞, 轩慧霞, 陈海良, 刘广臣, 李英乾, 管永, 管清宝, 游正伟. 多重动态杂化交联构建聚脲氨酯弹性体及其自愈合与可回收再加工性能. *高分子学报*, **2025**, 56(11), 1976–1986.
- 8 吴佳妮, 王岳鹏, 钱博, 吴泽凯, 苏基林, 王艺涵, 李闯, 游正伟. 基于受阻脲键的可重加工聚(脲-氨酯)弹性体的无溶剂制备、性能及应用研究. *高分子学报*, **2025**, 56(9), 1546–1556.
- 9 Gao, L.; Lin, J. P.; Wang, L. Q.; Du, L. Machine learning-assisted design of advanced polymeric materials. *Acc. Mater. Res.*, **2024**, 5(5), 571–584.
- 10 Behera, P. K.; Raut, S. K.; Mondal, P.; Sarkar, S.; Singha, N. K. Self-healable polyurethane elastomer based on dual dynamic covalent chemistry using Diels-Alder “click” and disulfide metathesis reactions. *ACS Appl. Polym. Mater.*, **2021**, 3(2), 847–856.
- 11 Heo, Y.; Sodano, H. A. Self-healing polyurethanes with shape recovery. *Adv. Funct. Mater.*, **2014**, 24(33), 5261–5268.
- 12 Tran, H.; Gurnani, R.; Kim, C.; Pilania, G.; Kwon, H. K.; Lively, R. P.; Ramprasad, R. Design of functional and sustainable polymers assisted by artificial intelligence. *Nat. Rev. Mater.*, **2024**, 9(12), 866–886.
- 13 Zhong, X. T.; Gallagher, B.; Liu, S. S.; Kailkhura, B.; Hiszpanski, A.; Han, T. Y. Explainable machine learning in materials science. *npj Comput. Mater.*, **2022**, 8, 204.
- 14 Ge, W.; De Silva, R.; Fan, Y. N.; Sisson, S. A.; Stenzel, M. H. Machine learning in polymer research. *Adv. Mater.*, **2025**, 37(11), 2413695.
- 15 Ding, F.; Liu, L. Y.; Liu, T. L.; Li, Y. Q.; Li, J. P.; Sun, Z. Y. Predicting the mechanical properties of polyurethane elastomers using machine learning. *Chinese J. Polym. Sci.*, **2023**, 41(3), 422–431.
- 16 Xu, P. C.; Ji, X. B.; Li, M. J.; Lu, W. C. Small data machine learning in materials science. *npj Comput. Mater.*, **2023**, 9, 42.
- 17 Han, Y. Z.; Du, W. T.; Zhang, Y. L.; Qiu, C.; Law, M.; Zhao, Y.; Tang, B. Z.; Wang, Y.; Yang, J. L. Programming interfacial polymerization: machine learning unveils quantitative rational design rules for microcapsules and beyond. *Adv. Mater.*, **2026**, 38(12), e17708.
- 18 Zhao, W. L.; Xu, X. Y.; Lan, H. X.; Wang, L. Q.; Lin, J. P.; Du, L.; Zhang, C. Y.; Tian, X. H. Designing multicomponent thermosetting resins through machine learning and high-throughput screening. *Macromolecules*, **2025**, 58(1), 744–753.
- 19 Deringer, V. L.; Bartók, A. P.; Bernstein, N.; Wilkins, D. M.; Ceriotti, M.; Csányi, G. Gaussian process regression for materials and molecules. *Chem. Rev.*, **2021**, 121(16), 10073–10141.
- 20 Oviedo, F.; Ferres, J. L.; Buonassisi, T.; Butler, K. T. Interpretable and explainable machine learning for materials science and chemistry. *Acc. Mater. Res.*, **2022**, 3(6), 597–607.
- 21 Wu, Y. H.; Wang, C.; Shen, X. T.; Zhang, T. Y.; Zhang, P.; Ji, J. Periodicity-aware deep learning for polymers. *Nat. Comput. Sci.*, **2025**, 5(12), 1214–1226.
- 22 Xie, C. H.; Qiu, H. K.; Liu, L.; You, Y.; Li, H. F.; Li, Y. Q.; Sun, Z. Y.; Lin, J. P.; An, L. J. Machine learning approaches in polymer science: progress and fundamental for a new paradigm. *SmartMat*, **2025**, 6, e1320.
- 23 柏康娜, 谢椿辉, 刘文涛, 犹阳, 李云琦. 聚氨酯类玻璃体应力松弛活化能的大数据解析. *高分子学报*, **2025**, 56(9), 1621–1632.
- 24 Li, R.; Lv, Y. J.; Xie, C. H.; Liu, L.; Ao, Q. L.; Li, Z.; Li, C. Y.; Li, Y. Q. Explore thermal and mechanical properties of biobased polyurethane elastomers through machine learning models. *Macromol. Rapid Commun.*, **2026**, 47(11), e00963.
- 25 Liu, L.; Li, R.; Xie, C. H.; You, Y.; Chen, Q.; Xie, H. B.; Qin, M. M.; Li, Y. Q. A big data approach to explore core properties of waterborne polyurethane coatings. *Prog. Org. Coat.*, **2026**, 211, 109739.

- 26 Ao, Q. L.; Xie, C. H.; You, Y.; Xie, H. B.; Qin, M. M.; Li, Y. Q. Data-driven models to explore regulatable variables for the density, thermal conductivity and compressive strength of rigid polyurethane foams. *Polymer*, **2026**, 346, 129592.
- 27 Mattia, J.; Painter, P. A comparison of hydrogen bonding and order in a polyurethane and poly(urethane-urea) and their blends with poly(ethylene glycol). *Macromolecules*, **2007**, 40(5), 1546–1554.
- 28 He, Y.; Xie, D. L.; Zhang, X. Y. The structure, microphase-separated morphology, and property of polyurethanes and polyureas. *J. Mater. Sci.*, **2014**, 49(21), 7339–7352.
- 29 Gandini, A. The furan/maleimide Diels-Alder reaction: a versatile click-unclick tool in macromolecular synthesis. *Prog. Polym. Sci.*, **2013**, 38(1), 1–29.
- 30 Kim, B. K.; Lee, S. Y.; Xu, M. Polyurethanes having shape memory effects. *Polymer*, **1996**, 37(26), 5781–5793.
- 31 Menon, A.; Thompson-Colón, J. A.; Washburn, N. R. Hierarchical machine learning model for mechanical property predictions of polyurethane elastomers from small datasets. *Front. Mater.*, **2019**, 6, 87.
- 32 De Gennes, P. G. Reptation of a polymer chain in the presence of fixed obstacles. *J. Chem. Phys.*, **1971**, 55(2), 572–579.
- 33 刘伦洋, 丁芳, 李云琦. 高分子材料大数据研究: 共性基础、进展及挑战. *高分子学报*, **2022**, 53(6), 564–580.
- 34 Scott, G. Properties of polymers. their correlation with chemical structure; their numerical estimation and prediction from additive group contributions. *Endeavour*, **1992**, 16(2), 97–98.
- 35 Fang, S. K.; Zeng, W. T.; Liu, L.; Lyu, Y.; Dong, F. W.; Wan, L.; Liu, Y. J.; Du, A. H. Shifting the mechanical strength-healing efficiency trade-off of polyurethanes by incorporating boronic ester bonds and hydrogen bonding. *Eur. Polym. J.*, **2024**, 206, 112776.
- 36 Li, Y. Q.; Liu, L. Y.; Chen, W. D.; An, L. J. Materials genome: research progress, challenges and outlook. *Sci. Sin.-Chim.*, **2018**, 48(3), 243–255.
- 37 Blackwell, J.; Gardner, K. H. Structure of the hard segments in polyurethane elastomers. *Polymer*, **1979**, 20(1), 13–17.
- 38 Sonnenschein, M. F.; Lysenko, Z.; Brune, D. A.; Wendt, B. L.; Schrock, A. K. Enhancing polyurethane properties via soft segment crystallization. *Polymer*, **2005**, 46(23), 10158–10166.
- 39 Kuenneth, C.; Rajan, A. C.; Tran, H.; Chen, L. H.; Kim, C.; Ramprasad, R. Polymer informatics with multi-task learning. *Patterns*, **2021**, 2(4), 100238.
- 40 Yan, Z. H.; Huang, H. X.; Ding, G. Q.; Dong, S. S.; Jiang, K. J.; Li, Y. Y.; Liu, Y.; Jiang, Y.; Wang, S. F.; Hu, G. H.; Du, J.; Zhang, S. X.; Zhao, H. Research progress of high-strength self-healing polymer materials: balance between mechanical strength and self-healing efficiency. *Chem. Eng. J.*, **2025**, 518, 164609.
- 41 Zhou, Y.; Li, L.; Han, Z. B.; Li, Q.; He, J. L.; Wang, Q. Self-healing polymers for electronics and energy devices. *Chem. Rev.*, **2023**, 123(2), 558–612.
- 42 Li, Y.; Zhou, M.; Wang, R. F.; Han, H. C.; Huang, Z.; Wang, J. Self-healing polyurethane elastomers: an essential review and prospects for future research. *Eur. Polym. J.*, **2024**, 214, 113159.
- 43 Liu, C.; Kelley, S. O.; Wang, Z. J. Self-healing materials for bioelectronic devices. *Adv. Mater.*, **2024**, 36(35), 2401219.
- 44 Huang, Q.; Li, Y. D.; Zhu, L.; Zhao, Q. B.; Yu, W. J. Unified multimodal multidomain polymer representation for property prediction. *npj Comput. Mater.*, **2025**, 11, 153.
- 45 Chen, T. Q.; Guestrin, C. XGBoost: a scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco California: ACM, **2016**. 785–794.
- 46 Gurnani, R.; Kuenneth, C.; Toland, A.; Ramprasad, R. Polymer informatics at scale with multitask graph neural networks. *Chem. Mater.*, **2023**, 35(4), 1560–1567.
- 47 Malashin, I.; Tynchenko, V.; Gantimurov, A.; Nelyub, V.; Borodulin, A. Boosting-based machine learning applications in polymer science: a review. *Polymers*, **2025**, 17(4), 499.
- 48 Lundberg, S.; Lee, S. I. A Unified approach to interpreting model predictions. *arXiv*, **2017**.
- 49 Hu, W. L.; Jing, E. Z.; Qiu, H. K.; Sun, Z. Y. Discovering polyimides and their composites with targeted mechanical properties through explainable machine learning. *J. Mater. Inform.*, **2025**, 5(1), 1–15.
- 50 Dalal, R. J.; Oviedo, F.; Leyden, M. C.; Reineke, T. M. Polymer design via SHAP and Bayesian machine learning optimizes pDNA and CRISPR ribonucleoprotein delivery. *Chem. Sci.*, **2024**, 15(19), 7219–7228.
- 51 Wang, H. R.; Cao, L.; Wang, X. L.; Lang, X. R.; Cong, W. W.; Han, L.; Zhang, H. Y.; Zhou, H. B.; Sun, J. J.; Zong, C. Z. Effects of isocyanate structure on the properties of polyurethane: synthesis, performance, and self-healing characteristics. *Polymers*, **2024**, 16(21), 3045.

- 52 Hao, Y. J.; Zhu, G. M. The latest advances in mechanically robust self-healing polyurea based on dynamic chemistry. *Adv. Sci.*, **2025**, 12(19), 2414788.
- 53 Ma, B. R.; Finan, N. J.; Jany, D.; Deagen, M. E.; Schadler, L. S.; Brinson, L. C. Machine-learning-assisted understanding of polymer nanocomposites composition-property relationship: a case study of NanoMine database. *Macromolecules*, **2023**, 56(11), 3945–3953.
- 54 Wang, W. W.; Chen, H. X.; Dai, Q. L.; Zhao, D.; Zhou, Y.; Wang, L. J.; Zeng, D. L. Thermally healable PTMG-based polyurethane elastomer with robust mechanical properties and high healing efficiency. *Smart Mater. Struct.*, **2019**, 28(1), 015008.
- 55 Ha, Y. M.; Kim, Y. O.; Ahn, S.; Lee, S. K.; Lee, J. S.; Park, M.; Chung, J. W.; Jung, Y. C. Robust and stretchable self-healing polyurethane based on polycarbonate diol with different soft-segment molecular weight for flexible devices. *Eur. Polym. J.*, **2019**, 118, 36–44.

## Research Article

## A Literature Data-driven Machine Learning Study of Structure–Property Relationship in Diels–Alder Polyurethane Elastomers Based on Hierarchical Domain-knowledge Features

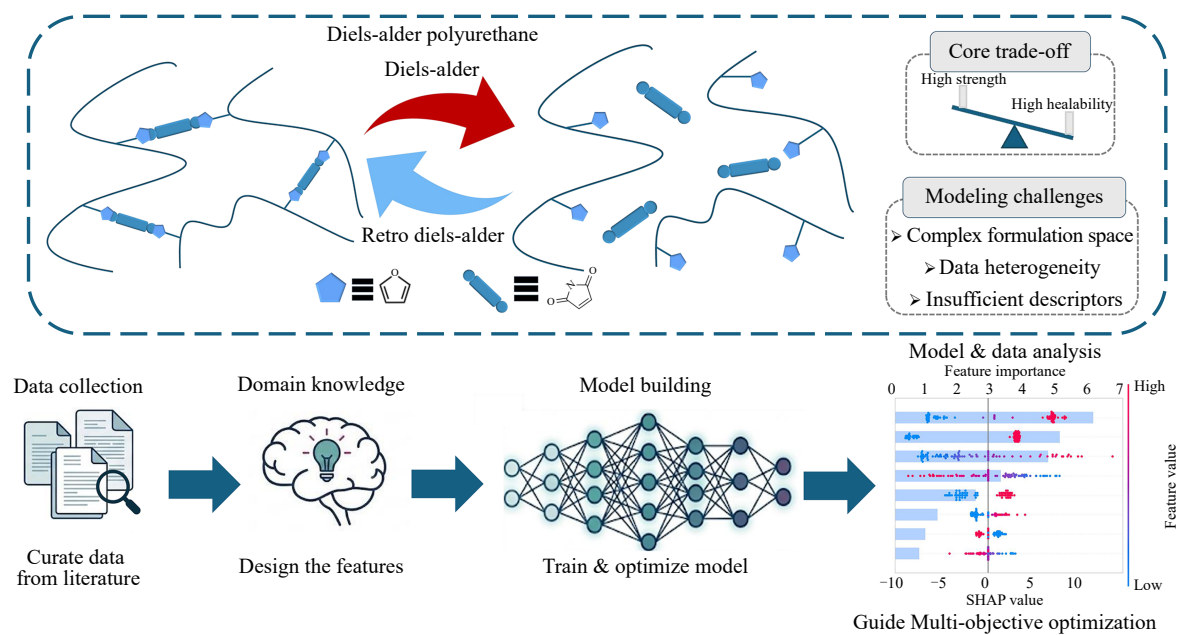
Jun-jie Yuanli<sup>†</sup>, Xiao-tian Chen<sup>†</sup>, Zheng-wei You<sup>\*</sup>

(State Key Laboratory of Advanced Fiber Materials, College of Materials Science and Engineering,  
Donghua University, Shanghai 201620)

**Abstract** Diels–Alder (DA) dynamic covalent polyurethane (DAPU) elastomers are highly regarded in the field of self-healing materials due to their excellent mechanical properties and thermo-reversible healing capabilities. In this study, we employed a literature data-driven approach to investigate the structure-property relationships governing the tensile strength, elongation at break, and healing efficiency of DAPU elastomers. By systematically screening the formulation and performance data from the relevant literature, we constructed a comprehensive feature space spanning the compositional content, chemical characteristics, and structural mechanisms. Three machine learning models, namely Support Vector Regression (SVR), XGBoost, and Gaussian Process Regression (GPR), were established to quantitatively analyze these properties, yielding  $R^2$  values of 0.76, 0.66, and 0.81, respectively. The results demonstrate that the proposed feature system effectively captures the relationships between the key structural variables and properties of DAPU elastomers. Subsequent SHapley Additive exPlanations (SHAP) analysis revealed that soft-segment crystallinity and soft-segment molecular weight were the most critical parameters governing mechanical strength and healing efficiency, whereas DA functional monomer content was the primary determinant of ductility. Furthermore, complex non-monotonic correlations were identified among the three properties. This work provides a reference for understanding DAPU structure-property relationships and guiding rational-formulation design.

<sup>†</sup> These authors contributed equally to this work.

<sup>\*</sup> Corresponding author: Zheng-wei You, E-mail: youzw@dhu.edu.cn



**Keywords** Polyurethane; Diels-Alder dynamic bonds; Self-healing elastomer; Machine Learning; Structure-property relationships; Literature data-driven