arXiv:2603.15686v1 [physics.chem-ph] 15 Mar 2026

所有有机化学品的生命周期评价

Shaohan Chen[1,2],Tim Langhorst[1],Julian Nöhl[1], Christopher Oberschelp[2,3], Martin Pillich[1], Johannes Schilling[1,*], 以及 André Bardow[1,2,*]

[1]能源与过程系统工程实验室（EPSE），苏黎世联邦理工学院，8092 苏黎世，瑞士[2] NCCR Catalysis，瑞士

[3]生态系统设计教席（ESD），苏黎世联邦理工学院，8092 苏黎世，瑞士[*]如需联系，请发邮件至： abardow@ethz.ch 和jschilling@ethz.ch

2026年3月18日

摘要

化学品渗透于现代社会的方方面面，但其生产引发了重大的可持续性问题。实现可持续的化学工业，需要详细的生命周期评价（LCA）；然而，由于现有生命周期清单（LCI）数据库仅涵盖了极少一部分交易中的化学品，加之数据覆盖有限、部分不一致且不透明，当前的评价仍面临诸多未知因素。在此，我们介绍 Chemical RetrosYnthesiS for Transparent Assessment of Life-cycles (CRYSTAL) 框架。该框架利用逆合成和机器学习生成的门到门清单，根据有机化学品的分子结构自动生成一致且透明的生命周期清单数据。借助 CRYSTAL 的预测能力，我们建立了一个涵盖超过70,000 种有机化学品的一致性数据库，其中包含超过110,000 份透明的生命周期清单数据集，量化了原料和能源需求，以及相关的辅助材料、生物圈流和废物流。基于这一综合数据库，我们确定了 50 个关键环境热点，它们在多个环境类别中驱动着有机化学品生产造成的高影响；我们还找出了对下游化学品生产至关重要的枢纽化学品。CRYSTAL 框架提供了全面的数据基础，可为有针对性的工程和政策干预提供系统指导。其透明、模块化的特性旨在推动化学品生命周期评价摆脱对“未知的未知”的依赖，转而建立一幅可通过协作不断完善的“已知的未知”图景。

1 引言

化学品及其衍生物存在于95% 的社会制成品中 [1]。其生产是化石资源消耗量最大的工业领域，也是全球第三大二氧化碳排放源$_2$ [2], ，并且是造成空气和水污染、严重威胁人类健康的主要因素 [2, 3]。因此，实现全球气候目标，要求化学工业迅速而稳健地转型，实现可持续发展。这一转型的进展取决于优先处理环境热点——即最有可能减少影响的工艺和产品——并替代令人担忧的化学品；欧盟委员会的安全与可持续设计倡议也倡导这一做法 [4, 5]。然而，这些环境热点在很大程度上仍然未知。

挑战在于规模和复杂性：40,000 至 60,000 种化学品在全球贸易中流通 [6], 并通过错综复杂的价值链相互关联。要有效减少由此产生的环境影响，需要开展综合系统分析，将上游价值链的潜在改进与下游生产的影响联系起来。此类综合系统分析要求全面、透明且一致地量化环境影响 [7] ，并以所有化学品生产流程的详细生命周期清单（LCI）数据（即相应的物质和能量流）为依据。然而，目前交易的大多数化学品都没有此类生命周期清单数据；在成熟的生命周期评价数据库中，只有不到 2,000 种化学品具备这类数据 [8, 9] （见图 1a）。扩充数据库需要针对每种缺失的化学品耗费大量时间进行人工数据收集，这使得大规模开展综合系统分析难以实现。这一瓶颈阻碍了研发工作聚焦于能够减少环境影响的化学品生产流程，从而延缓了化学工业向可持续发展转型。

将生命周期评价的覆盖范围扩展至整个化学品行业，促使人们运用机器学习来预测化学品生产的环境影响 [10, 11, 12, 13]。这些方法有望实现规模化应用，但受限于生命周期评价训练数据的匮乏，而且往往作为黑箱运行，难以解释底层反应路径或生命周期清单数据。SemaNet 等工作 [14] 通过从文本工艺描述中推断缺失的生命周期清单流来弥补数据缺口，但并未明确捕捉化学品生产的底层反应路径。第一性原理方法（如基于化学计量的方法） [15], 提供了更透明的替代方案，但需要详细了解底层反应路径，而从文献中提取这些知识仍然耗费大量人力和时间。

在此，我们介绍用于透明生命周期评价的化学逆合成（Chemical RetrosYnthesiS for Transparent Assessment of Life-cycles，CRYSTAL）。这是一种完全预测性、按生产路径解析的有机化学品生产生命周期评价框架，仅需输入目标化学品的分子结构。CRYSTAL 可自动生成透明且一致的生命周期清单数据，并确定环境最优的生产路径，从而能够对有机化学品开展大规模环境影响评估。利用 CRYSTAL，我们绘制了全球化学反应网络，识别出关键环境热点

，比较了多个影响类别相对于气候变化的表现，并突出显示了若加以改进即可带来最大可持续性收益的关键过程。

2 CRYSTAL框架：有机化学品生命周期清单的自动生成

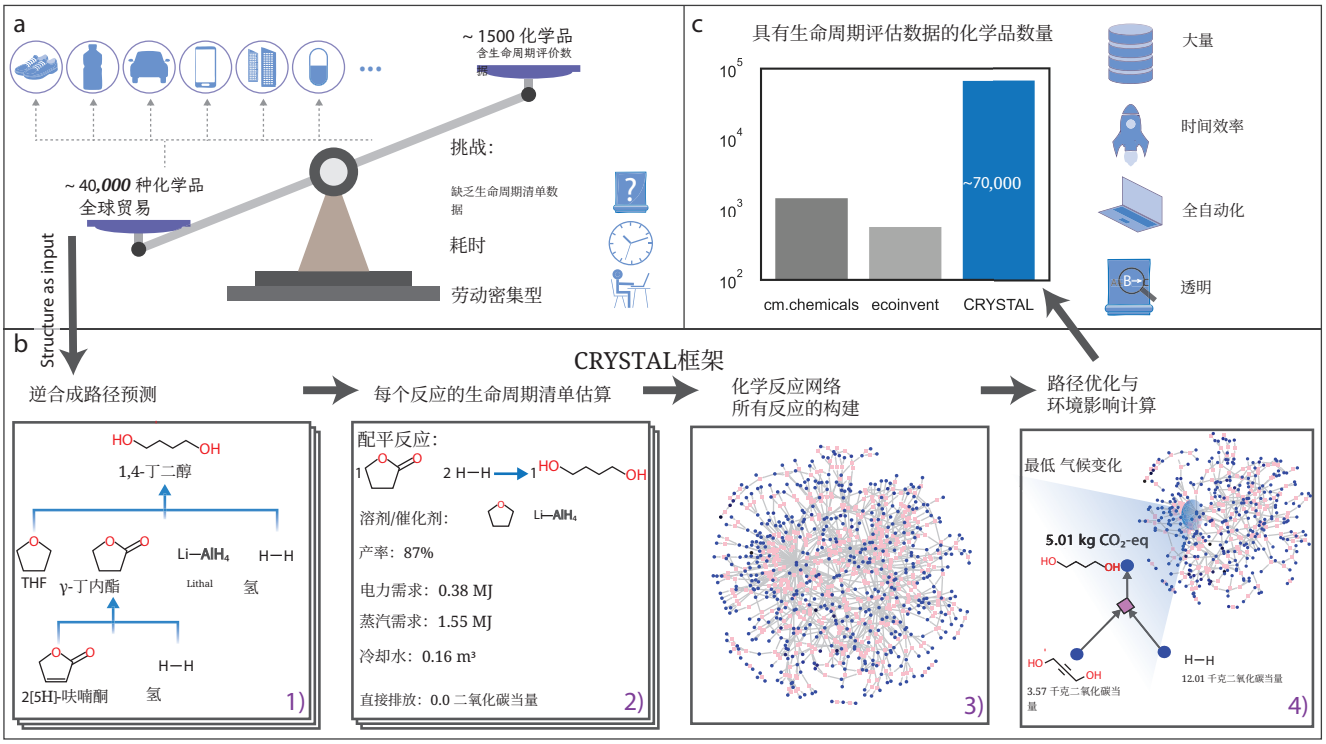


图1： CRYSTAL框架可自动、透明地预测有机化学品生产从摇篮到工厂大门的生命周期清单数据，并识别环境表现最优的生产路径。 CRYSTAL在高效生成生命周期清单数据与评估大量贸易化学品环境影响的需求之间取得平衡，弥补生命周期评价中的关键数据缺口。基于目标化学品的分子结构，我们 1）预测逆合成路径；2）估算每个反应步骤的生命周期清单数据，即物料需求、能源需求、废物流和直接排放；3）构建化学反应网络（CRN），以在各反应之间传递信息；4）根据预测的生命周期清单数据计算生命周期影响评价（LCIA）得分，并据此优化化学反应网络，结合用户指定的环境目标，识别环境表现最优的生产路径。该框架可灵活适配各种生命周期影响评价方法和背景数据库。

CRYSTAL框架（图1b）整合化学、工程学和优化领域的知识，以系统且一致的方式量化有机化学品从摇篮到工厂大门的优化环境影响（详情见补充文本）。对于每种新的目标化学品，CRYSTAL首先利用机器学习驱动的逆合成生成可行的合成路线 [16], ，由此确定化学反应的反应物和辅助物。这些合成路线沿价值链递归地反向生成，直至所有前体均可在底层生命周期评价数据库（如 ecoinvent [17, 18, 8]

或 cm.chemicals [9, 19]）中找到。由于生产有机化学品时偶尔需要无机化合物作为反应物或辅助材料，我们将文献报道的无机反应路径纳入考量，以补充逆合成工具，为底层生命周期评价数据库中未收录的关键无机前体提供相关路径。随后，通过优化算法对识别出的反应方程式进行配平并识别副产物，以估算原料需求、副产物产量以及废物流。接下来，利用基于工业过程数据训练的决策树预测反应收率、能源需求和辅助材料需求 [20]。废物和副产物的处理由层级规则模型指导，该模型整合了行业最佳实践（详情见补充文本）和废物焚烧模型 [21, 22, 23]。由此即可获得所有必需的生命周期清单数据，并计算目标化学品相应的生命周期影响评价（LCIA）得分。CRYSTAL的模块化结构便于高效替换各个模块（例如，底层生命周期评价数据库或能源需求估算方法）。

由于逆合成工具并非旨在预测环境效益最优的反应路径，而且通过共享中间体可能找到改进后的路径，因此我们将所有已识别的反应整合到一个统一的化学反应网络（CRN）中。我们使用基于图论的算法对该化学反应网络进行优化，以传播生命周期清单数据信息并识别环境效益更优的路径。如有需要，还可通过纳入与工业相关的反应，灵活扩展该化学反应网络。

通过采用这一全面的方法，CRYSTAL 可自动识别最优、透明的生产路径，并同时提供相应的生命周期清单数据，一致地量化相应的环境影响。我们利用 CRYSTAL 对从多个监管和工业来源中识别出的具有市场相关性的化学品进行高通量筛选，这些来源包括 REACH数据库 [24], DIPPR数据库 [25], 经合组织数据库 [26], 美国环境保护署化学品数据报告数据库 [27], 斯堪的纳维亚北欧SPIN数据库 [28] 以及日本经济产业省数据库 [29] （详情见补充文本）。这项筛选在数小时内即可为约 70000 种目标化学品及其中间体生成一致的生命周期清单数据和环境效益优化的生产路径。由此得到的生命周期评价数据库 40 倍于现有最大的生命周期评价数据库 [8, 9] （图 1c），并可利用 CRYSTAL 持续扩展，以弥补生命周期评价数据的关键缺口并支持未来研究。

3 验证 基于先进生命周期评价数据库的 CRYSTAL s

与现有的机器学习方法仅预测化学品汇总环境影响不同，CRYSTAL 以透明的方式预测底层生命周期清单数据，并以此作为计算相应环境影响的基础。由于此类生命周期评价数据几乎不存在，验证 CRYSTAL 这类方法便面临一项特殊挑战。我们通过比较其预测的环境

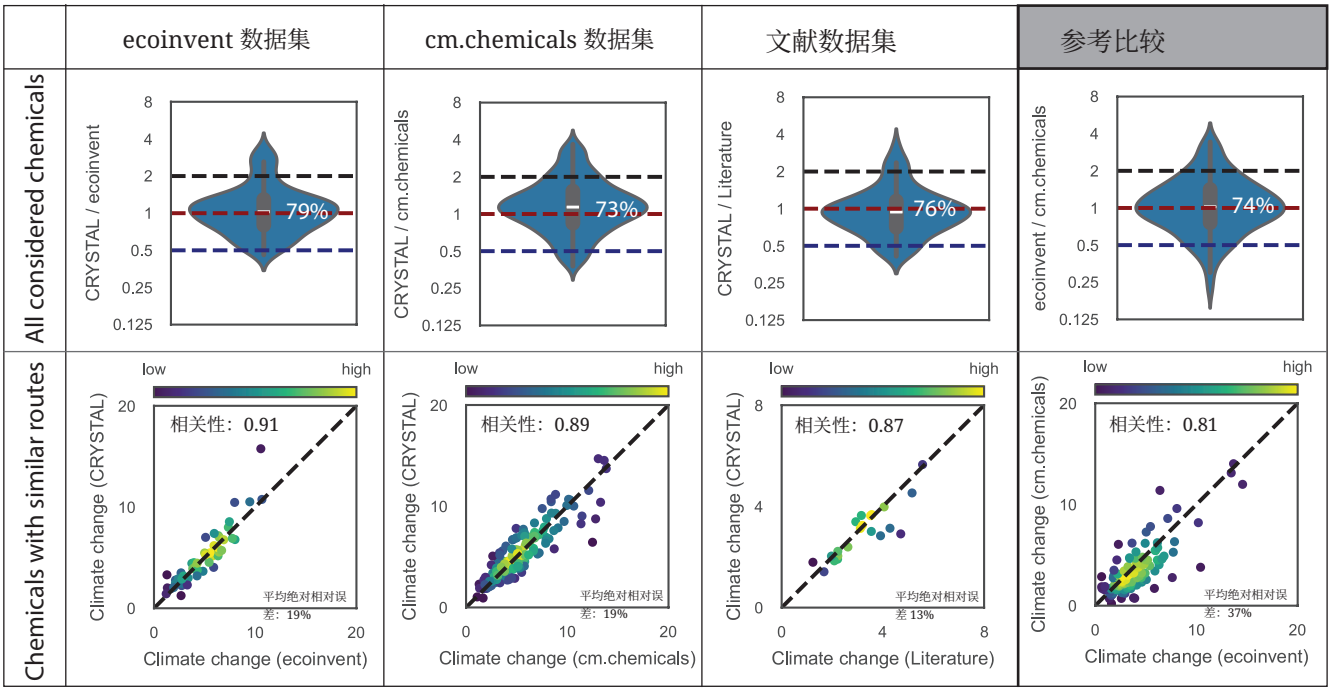


图2： CRYSTAL在三个最先进的生命周期评价数据库中的气候变化影响表现。前三列展示了CRYSTAL框架在ecoinvent中的验证结果 [17, 18, 8] (257 种有机化学品（3.10版，截断法），采用ReCiPe 2016 v1.03中点（H）方法进行评估）、cm.chemicals [9, 19] (746 种基于IPCC 2021全球变暖潜势（GWP100）的有机化学品），以及一个文献数据库 [15] (143 种采用ecoinvent 3.5版截断法、基于ReCiPe v1.13中点（H）方法进行评估的有机化学品）的验证结果；验证以相应数据库为依据，并采用留一法。“参考比较”列针对232种化学品，对比ecoinvent与cm.chemicals数据库；这些化学品是两个数据库中均有记录且具有可用SMILES表示的化学品的完整交集，可作为评估CRYSTAL性能的基准。前三列的小提琴图汇总了CRYSTAL在各数据库中的整体表现，显示基于预测值与参考值之比，预测结果落在可接受准确度范围内的气候变化(单位：千克CO_2-当量/千克目标化学品）的百分比。虚线表示由成本工程促进协会（AACE）准确度界限所定义的可接受范围（即相应最先进数据库所定义环境影响的50% 至200% ）。小提琴图展示第5至第95百分位数之间的分布，以突出这一范围。前三列的一致性图展示了当预测路线的最后一个反应步骤与参考数据库中的反应步骤相同时，CRYSTAL的皮尔逊相关系数和平均绝对相对误差（MARE）。

与三个最先进的生命周期评价数据库 [8, 9, 15], 中的参考值进行影响比较，此处以气候变化影响为例（图2）。为确保比较的一致性，我们将每个最先进数据库作为 CRYSTAL 的底层参考，并采用留一法（LOO）（详见补充文本），协调不同数据集中的背景过程和模型假设。

超过73% 的 CRYSTAL 预测值与最先进参考影响值的差异在2 倍以内；根据成本工程师促进协会国际组织的成本工程标准，这一范围被定义为可接受范围（第 5 类：50% 至200%） [30]。我们的结果表明，CRYSTAL 的偏差与现有最先进生命周期评价数据库之间观察到的差异处于同一水平（图2中“参考比较”列）。数据库之间的此类差异很大程度上源于上游背景过程和模型假设的不同（例如，全球化学品生产的市场结构） [31]。值得注意的是，所考虑的 cm.chemicals 数据库版本 [9, 19] 采用 ecoinvent 作为其背景数据库。因此，完全独立的生命周期评价数据库之间的比较预计会呈现更大的差异。

CRYSTAL 与参考数据库之间剩余偏差的一个来源，是合成同一种化学品时采用了不同的生产路径。由于实际中会采用不同的生产路径，因此这类差异不能先验地视为错误。当最后一个预测的反应步骤与参考路径一致时，CRYSTAL 表现出很高的预测准确性，皮尔逊相关系数达到0.9 ，平均绝对相对误差为19%，其中96% 和100 % 的预测结果分别落在 ecoinvent 的可接受范围内 [8, 17, 18] 和 cm.chemicals [9, 19]，（图2）。当预测只涉及一个反应步骤时，CRYSTAL 的预测结果与 ecoinvent[8, 17, 18] 和 cm.chemicals [9] 数据库最为接近，分别有81% 和79% 的预测结果落在可接受范围内（图 S37 和 S38）。随着预测的反应步骤数量增加，CRYSTAL 的预测结果与参考值之间的偏差也随之增大，这是由于预测路径与参考路径之间的差异不断累积。在 LOO 方法中，通常需要预测多个反应步骤就表明预测路径与参考路径存在差异，因为原则上，所有前体都应能在底层生命周期评价数据库中找到。

CRYSTAL 的泛化能力并不限于气候变化影响，这一点已通过对所有环境影响类别的验证得到证明（图 S28–S36）。具体而言，当预测路径的最后一个反应步骤与参考路径一致时，CRYSTAL 对能源资源：非可再生、化石的预测中有 87% 落在规定的可接受范围内，皮尔逊相关系数达到 0.89。

为进一步验证 CRYSTAL 在结构和功能更复杂的化学品（如药品）上的性能，我们将 CRYSTAL 与两个已发表数据库 [32, 33] 以及工业合作伙伴提供的一个工业数据集进行对比（图 S39、S41 和 S43）。在预测相同的最后一个反应步骤时，CRYSTAL 成功再现了参考数据中环境影响的整体趋势及相对大小，皮尔逊相关系数分别达到 0.95 和 0.98（采用对数值以校正尺度差异

（跨化学品）在两个已发表数据库上 [32, 33]。对于工业数据集，尽管由于保密原因缺少反应路径的详细信息，CRYSTAL 的整体表现仍达到皮尔逊相关系数0.81 。

尽管生命周期评价缺乏可用于验证的确定性真实值 [34]，我们的结果表明，CRYSTAL 与既有基准高度一致。我们承认，CRYSTAL 对环境影响给出了较为乐观的估计，因为该框架系统地识别出当前已建模的最有利化学路线。尽管如此，通过将生命周期评价的覆盖范围扩展至数以万计的化学品，CRYSTAL 可作为一个透明的起点，以弥补关键数据缺口、激励共享一手工业数据，并加强整个化学工业的生命周期评价研究。

4 有机化学品生产中超越气候变化的环境热点

CRYSTAL 能够系统地识别化学品生产价值链中的环境热点（详见补充文本），包括原料、辅助材料、废物处理和能源需求。由于当前的评估实践主要聚焦于气候变化[35]，我们识别出对其他环境影响贡献不成比例的相对环境热点。在 CRYSTAL数据库中约70000 种化学品中，1616 种化学品在其气候变化优化路径中表现出不成比例的高人体致癌毒性（图3 上排）。追溯这些化学品的完整生产价值链可发现，两种上游反应物是主要贡献者（详见补充文本）：三氧化铬（CAS 登记号 1333-82-0，来自 ecoinvent 的 LCIA 数据——1451 种下游产品）、蒽醌（CAS 登记号 84-65-1，来自 ecoinvent 的 LCIA 数据——62 种下游产品）。其余具有高人体致癌毒性的化学品主要源于含铬废物处理。

与热点三氧化铬和蒽醌相关的较高人体致癌毒性，既源于它们的生产，也源于上游过程中对（含铬）废物的处理，ecoinvent [36]对此已有报告。CRYSTAL 中实施的废物处理模型（详情见补充文本）假设，无价值副产品（包括含铬废物）会被焚烧[37, 38, 39, 40] ，遵循 Doka 模型 [21, 22]，，从而导致六价铬进一步排放到地下水、河流和大气中 [21]。然而，日益加强的监管压力 [41, 42] 正推动化学工业减少六价铬排放。一项有效的减缓策略是将六价铬转化为毒性较低且更稳定的三价铬 [43]。在我们的模拟中，我们在三氧化铬和蒽醌的生产价值链内将六价铬转化为三价铬（详情见补充文本），

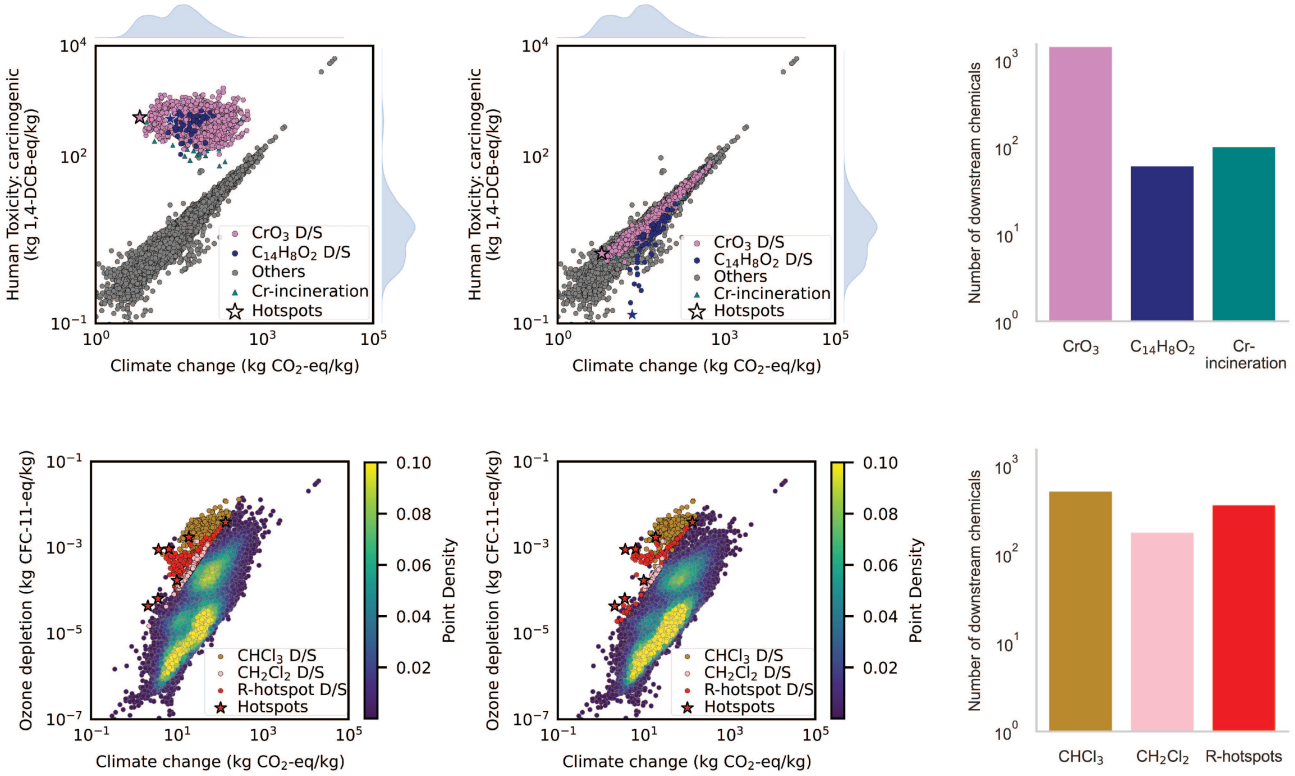


图3： 定位各影响类别中相对于气候变化的环境热点，以支持化学品制造商、监管机构和 生命周期评估数据提供方。 上排： 气候变化 与 人体毒性：致癌 针对气候变化 影响优化的 路径。三氧化铬和蒽醌，以及含铬废物焚烧，被标记为 人体毒性：致癌方面的相对环境热点，同时还包括其下游化学品（左）。对这两种热点反应物的六价铬排放以及含铬废物处理进行监管，可有效降低所有下游化学品的 人体毒性：致癌 影响（中）。与这些热点相关的下游化学品数量（D/S）（右）。下排： 气候变化 与 臭氧消耗 针对（左） 气候变化 和（中） 臭氧消耗 影响优化的路径。气候变化优化路径中的相对环境热点及其下游化学品一并标出（即五种上游反应物（R类热点）和两种溶剂：氯仿和二氯甲烷）。为突出点密度的变化，仅标出臭氧消耗影响比例异常高的下游化学品（中）。与这些热点相关的下游化学品数量（右）。采用 ecoinvent 3.10版（截断法）[17, 18, 8] 以及 LCIA 方法 ReCiPe 2016 v1.03、中点（H）作为基础生命周期评价数据库。

以及在含铬废物处理过程中采取这些措施，可有效消除这些热点及其下游化学品所表现出的异常高人体致癌毒性（图3，上排，中间）。我们的研究结果凸显了将六价铬还原措施确立为废物焚烧和填埋领域工业标准的重要性，并强调亟须确保将现行监管实践纳入生命周期清单数据库和废物处理模型 [21, 22]。

这些结果表明，在被确定为相对环境热点的化学品生产上游阶段实施针对性监管和工艺改进，可以显著降低人体毒性影响。其他环境影响类别的结果见图S46–S53。更广泛而言，我们的研究结果强调，有必要将最新的工业实践和监管进展纳入生命周期评价清单和模型，包括当前的CRYSTAL框架，以确保环境影响评估准确，并为可持续化学品设计和政策制定提供可靠指导。

针对臭氧消耗与气候变化之间的权衡，我们使用高斯混合聚类算法，沿着气候变化优化路径识别出五类化学品（图3 —下排，左侧；见扩展数据图1a及补充文本中的详细说明）。在臭氧消耗影响异常高的化学品中，我们确定了七种上游贡献因素，它们是臭氧消耗的主要驱动因素（图3 —下方，左侧；方法见补充文本）：两种溶剂，氯仿（CAS 登记号 67-66-3，数据来自 ecoinvent——525 种下游产品）和二氯甲烷（CAS 登记号 75-09-2，数据来自 ecoinvent——181 种下游产品）；以及五种反应物（数据均来自 ecoinvent——共计368 种下游产品）。这些贡献因素造成的臭氧消耗影响较高，原因在于生产过程中直接排放卤代烷烃或一氧化二氮，或将氯仿用作前体。

将目标从最小化气候变化影响转为最小化臭氧消耗后，27600 种化学品采用了替代生产路径，从而减轻了气候变化优化路径所带来的臭氧消耗影响。对于这些化学品，臭氧消耗影响可大幅降低50 %，而气候变化影响仅增加19%（图3 – 底部，中间）。五个聚类质心中的臭氧消耗影响均有所下降，其中初始影响最高的聚类降幅最大（见扩展数据图1b）。然而，对于某些化学品，臭氧消耗影响的缓解仍受化学反应网络中可用反应的限制。

我们的分析进一步印证了《蒙特利尔议定书》的目标[44] ，展示了减少或替代作为溶剂使用的氯仿和二氯甲烷，以减缓臭氧消耗所带来的潜在环境效益。正如已识别出的五种上游热点反应物所表明的那样，严格控制卤代烷烃和一氧化二氮的直接排放仍至关重要。随着逆合成工具 [16] 通过更新训练数据不断发展，以反映化学工业的持续进步，同时生成的生命周期清单数据也得到扩充，

因此，CRYSTAL框架将通过确定化学品生产中的可行干预措施来支持这项工作。

5 哪些化学品对提升可持续性最重要?

基于其底层化学反应网络，CRYSTAL 揭示了枢纽化学品——这些关键中间体的环境影响会波及化学反应网络中的大片区域。由于它们在连通性和物质流动中发挥着核心作用（见图 4a），改善这些枢纽化学品的可持续性能够将益处扩展至整个化学工业 [45]。在此，我们将枢纽化学品定义为满足以下条件的化学品：(i) 出现在至少 216 条经过环境优化的下游路径中（通过肘部法确定，见扩展数据图 2）；且 (ii) 会产生可测量的传播效应，即其环境影响降低 10 %，能够使超过 2%的环境影响降低，且受影响的下游化学品超过 10 种。我们将这些下游化学品称为强烈受影响的节点。

我们识别出 52 种枢纽化学品，它们在针对气候变化优化的化学反应网络中具有广泛的系统性影响（图 4b；详细信息见表S9）。在这些枢纽化学品中，四氢呋喃（THF；CAS登记号109-99-9）的影响尤为显著：它能强烈影响超过 17 500 种下游化学品（占其下游化学品总数的 68%），平均使气候变化影响降低 4.5% 。这种影响沿层级向下传播，最多跨越 18 个下游反应步骤（扩展数据图 3a），反映出四氢呋喃兼具原料和溶剂的双重作用。四氢呋喃作为反应介质被广泛使用 [46] 进一步凸显了它在有针对性地改进工艺、实现系统层面气候效益方面的战略重要性。

二甲基甲酰胺（DMF；CAS登记号68-12-2）也呈现出类似的模式，对 7000 种下游化学品产生显著影响，气候变化影响平均降低 4。0% ，这再次体现了它兼具原料和溶剂的双重功能（扩展数据图 3b）。

尽管甲醇和乙醇分别出现在超过 25 000 条和 19 900 条下游生产路径中，但就气候变化而言，受到显著影响的下游化学品仅占 5% 和 6% （扩展数据图 3 b、c）。这种有限的影响传递，反映了它们自身生产路径的气候变化影响相对较低：乙醇采用生物质发酵，甲醇采用天然气重整。在我们的化学反应网络（CRN）中，与化石基生产相比，生物基乙醇使 2047 种强受影响下游化学品的气候变化影响平均降低 6kg CO_2-eq，凸显了枢纽化学品转向生物基生产可带来的显著气候效益（图 4c）。然而，这一转变也伴随着环境权衡，尤其是土地和水资源使用影响增加（图S49和S51）。

在气候变化方面排名靠前的枢纽化学品，往往也会对其他环境影响类别产生显著影响（扩展数据图4；详情见补充文本）。然而，这种模式

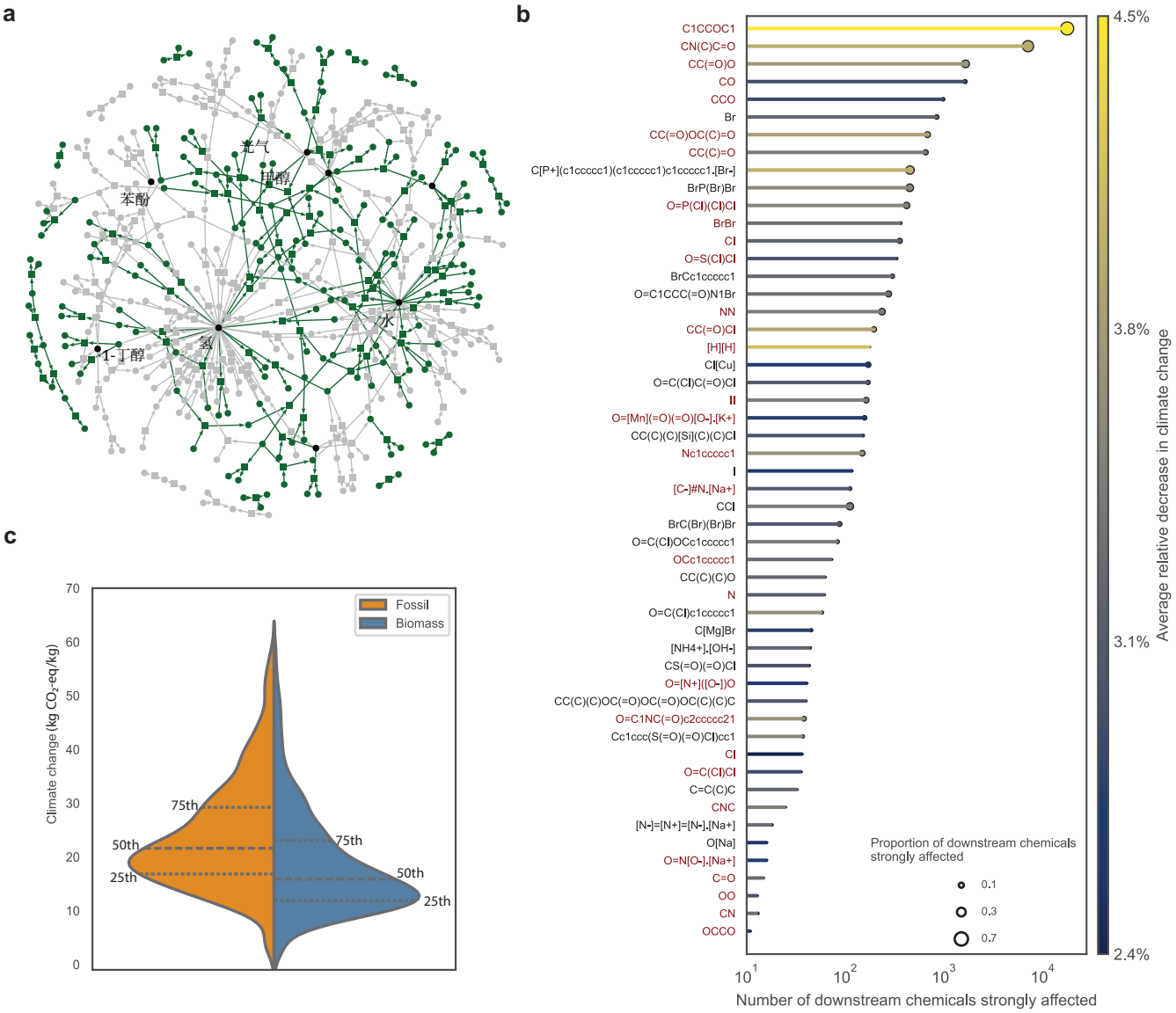


图 4：为提升化学工业的可持续性而确定的枢纽化学品，可用于指导过程优化策略的优先排序。a) 一个示例，展示了枢纽化学品在示意性化学反应网络中具有重要的结构作用。绿色部分表示所有环境影响类别下的联合最优反应。**b)** 按受强烈影响的下游化学品数量对 52 种枢纽化学品进行排序。受强烈影响的下游化学品是指：当某种枢纽化学品的影响降低 10% 时，其气候变化影响随之降低超过 2% 的化学品（详情见补充文本）。底层 ecoinvent数据库中已有的化学品以红色标出。**c)** 从化石基乙醇路径转向生物基乙醇路径，可显著降低 2000 多种受强影响下游产品的气候变化影响。

但并非总是如此：例如，硝酸因一氧化二氮排放而对臭氧消耗的影响波及超过 2800 种下游化学品，但对其他影响类别的影响有限。这些发现凸显了在气候变化之外考虑多个环境维度的重要性，因为某些枢纽化学品会产生特定类别的影响，而这些影响在侧重气候变化的评估中可能会被忽略。对不同环境影响类别下优化后的 化学反应网络进行比较发现，臭氧消耗和海洋富营养化与气候变化的相关性最低（见扩展数据图 5；以及补充文本中的分析），因此在分析化学品生产中的环境权衡时，应优先考虑这两项。

尽管受模型局限性影响，已识别的枢纽化学品可能无法反映现实世界中的所有影响，但结果表明，关注具有级联下游效应的枢纽化学品，能够为系统范围的可持续性改进提供参考；随着工业数据可用性的提高以及模型假设的不断完善，这一点尤为重要。

6 推动化学工业实现可持续发展

CRYSTAL 通过透明且一致地根据逆合成路径预测生命周期清单数据，标志着化学品生产环境影响评估的一项重大进步。CRYSTAL 自动为有机化学品提供生命周期清单，采用一致的假设、透明的反应路径和统一的详细信息，从而弥补了生命周期评价中的一项关键数据缺口。因此，我们相信，该数据库将对生命周期评价从业者大有裨益。不过，鉴于 G.E.P. Box 的名言“所有模型都是错的，但有些模型有用”，我们认为，透明性是 CRYSTAL 模型和数据库的关键特征：如果专家知道实际生产中某条反应路径、溶剂或产率有所不同，就可以替换生命周期清单中的相应组成部分。为最大限度地实现持续完善，CRYSTAL 被设计为一个开放的、以社区为导向的平台（采用开源代码和版本化数据集），欢迎研究人员、行业人士和数据提供者贡献内容。社区维护和结构化反馈机制将逐步提升预测质量。

尽管所识别的环境优化路径可能无法反映现实中经过经济优化的工业实践，因而不能直接用于环境报告，但 CRYSTAL 有助于发现提升有机化学品可持续性的机遇。通过指出可供重点优化的枢纽化学品和热点过程，CRYSTAL 为工业界、监管机构和生命周期评估数据提供方提供了切实可行的指导。

由于逆合成路径预测基于可获得的专利文献，CRYSTAL 还可以通过纳入基于生物基反应的路径[47, 48] 和塑料废物回收的路径 [49], 得到进一步增强，从而实现关键化学品中化石原料的替代，并带来级联式的系统范围环境效益（第 4 节和第 5 节）。纳入生产量将实现行业范围优化，以识别联合化学品生产中的协同效应

。将 CRYSTAL 与逆合成规划相结合，可以引导化学合成 [50]采用本质上更环保的路径，从而建立可持续化学创新的自我强化循环。

随着新型化学品发现步伐的加快，CRYSTAL能够从最早期阶段开始，提供快速、自动化且透明的可持续性评估。这一能力使化学家和工程师能够快速评估新兴路径，确保创新在不产生意外环境权衡的情况下推进，并促进形成兼具生产力与可持续性的化学工业。更重要的是，通过将社区反馈纳入模型和数据更新，CRYSTAL不仅定位为一种工具，也定位为一个协作生态系统，以持续改进生命周期清单预测及其与现实世界的相关性。

数据和材料可用性

支持本研究结论的所有数据均可在正文或补充材料中获取，但源自商业数据库的数据及商业数据库本身除外，因为这些内容无法再分发。完整的CRYSTAL框架及相应的CRYSTAL数据库将尽快公开。有关材料的通信和请求应发送至A.B。

补充材料

补充文本（包括其中提到的图S28–S39、图S41、图S43、图S46–S53和表S9）未包含在本arXiv预印本中，将在后续版本中提供。

参考文献

[1]国际化学品协会理事会（ICCA）与牛津经济研究院。《全球化学品行业：推动增长，应对世界的可持续发展挑战》。技术报告，ICCA / 牛津经济研究院（2019年）。网址 https://icca-chem.org/wp-content/uploads/2020/10/Catalyzing-Growth-and-Addressing-Our-Worlds-Sustainability-Challenges-Report.pdf。

[2] Pales, A. F. 与 Levi, P.。《石油化学品的未来：迈向更可持续的塑料和化肥》。技术报告，国际能源署（IEA），法国巴黎（2018）。网址 https://www.iea.org/reports/the-future-of-petrochemicals。

[3] Woodruff, T. J. 化石燃料衍生的内分泌干扰物的健康影响。《新英格兰医学杂志》**390**，922–933（2024年）。网址 https://doi.org/10.1056/NEJMra2300476。

[4]欧盟委员会，联合研究中心。安全与可持续设计。https://research-and-innovation.ec.europa.eu/research-area/industrial-research-and-innovation/chemicals-and-advanced-materials/safe-and-sustainable-design_en（2023年）。访问日期：2025-10-17。

[5] van der Waals, J. 等人 材料与化学品的安全设计：迈向地平线欧洲创新计划。报告，荷兰基础设施与水资源管理部（2019年）。URLhttps://doi.org/10.5281/zenodo.3254382。访问日期：2025年10月17日。

[6] 联合国环境署与国际化学品协会理事会（ICCA）。工业化学品健全管理的知识管理与信息共享。技术报告，联合国环境署（UNEP）/ICCA（2015年）。URL https://www.saicm.org/Portals/12/Documents/EPI/Knowledge_Information_Sharing_Study_UNEP_ICCA.pdf。访问日期：2025年10月17日。

[7] 国际标准化组织（ISO）。环境管理：生命周期评价——原则与框架。国际标准 ISO 14040:2006，ISO，日内瓦，瑞士（2006年）。

[8] ecoinvent协会。ecoinvent数据库第3版。 https://ecoinvent.org/database/ （2025年）。访问日期：2025-10-17。

[9] Carbon Minds GmbH。cm.chemicals 数据库：化学品和塑料的生命周期清单数据。 https://www.carbon-minds.com/products/data/carbon-footprint-and-lca-data/（2025年）。访问日期：2025-10- 17。

[10] 张 D.、王 Z.、Oberschelp C.、Bradford E. 和 Hellweg S. 化学品碳足迹增强型深度学习模型。《*ACS*可持续化学与工程》**12**，2700–2708（2024年）。URL https://doi.org/10.1021/acssuschemeng.3c07038。

[11] Kleinekorte, J. 等。 APPROPRIATE生命周期评价：针对新兴化学工艺的面向工艺、预测性影响评估方法。《*ACS*可持续化学与工程》**11**，9303–9319（2023年）。URL https://doi.org/10.1021/acssuschemeng.2c07682。

[12] Baxevanidis, P., Papadokonstantakis, S., Kokossis, A. & Marcoulaki, E. 基于基团贡献的生命周期评价模型，可用于计算机辅助分子设计应用中筛选环境友好型新型化学品。*AIChE*期刊 **68**, e17544 (2022). URL https://doi.org/10.1002/aic.17544.

[13] Gao, Q. 等 基于分子图和图神经网络的环境影响预测。《计算机与化学工程》**204**，109362（2026年）。网址 https://doi.org/10.1016/j.compchemeng.2025.109362。

[14] Chen, B. 等 SemaNet：连接文字与数字，预测生命周期评价中缺失的环境数据。《环境科学与技术》**59**，21131–21146（2025年）。网址 https://doi.org/10.1021/acs.est.5c07557。

[15] Langhorst, T., Winter, B., Roskosch, D. & Bardow, A. 基于化学计量的估算：新兴化学工艺的气候影响——方法基准测试与建议。《*ACS*可持续化学与工程》 **11**, 6600–6609 (2023). URL https://doi.org/10.1021/acssuschemeng.2c07624.

[16] Schwaller, P. 等 利用基于Transformer的模型和超图探索策略预测逆合成路径。《化学科学》 **11**, 3316–3325 (2020). URL https://doi.org/10.1039/C9SC05704H.

[17] Wernet, G. 等 ecoinvent数据库第3版（第一部分）：概述与方法论。《国际生命周期评价杂志》 **21**, 1218–1230 (2016). URL https://doi.org/10.1007/s11367-016-1087-8.

[18] Frischknecht, R. 等 ecoinvent数据库：概述与方法学框架。《国际生命周期评价杂志》**10**，3–9 (2005)。网址 https://doi.org/10.1065/lca2004.10.181.1。

[19]Stellner, L. 等 cm.chemicals 方法论。技术报告，Carbon Minds GmbH，德国科隆 (2023)。网址 https://www.carbon-minds.com/wp-content/uploads/2024/01/cmchemicals-methodology.pdf。2023年7月，版本 2.00。

[20] Langhorst, T., Winter, B., Tuchschmid, M., Roskosch, D. & Bardow, A. 从反应化学计量到生命周期评价：基于决策树的估算工具。《*ACS*环境*Au*》(2025年)。网址 https://doi.org/10.1021/acsenvironau.4c00065。

[21] Doka, G. 废物处理服务的生命周期清单。技术报告，ecoinvent，苏黎世，瑞士（2003年）。URL https://www.doka.ch/13_I_WasteTreatmentGeneral.pdf。访问日期：2025年10月17日。

[22] Doka, G. 废物处理服务生命周期清单的更新——第二部分：废物焚烧。技术报告，Doka生命周期评价，苏黎世，苏黎世，瑞士（2013年）。URL http://www.lcaforum.ch/inventories/Hintergrund/Doka_2013-ecoinvent_MSWI_updateLCI.pdf。访问日期：2025年10月17日。

[23] Meys, R. 等。 通过循环碳经济实现塑料温室气体净零排放。《科学》**374**，71–76（2021年）。URL https://doi.org/10.1126/science.abg9853。

[24]欧洲化学品管理局 (ECHA)。REACH——化学品注册、评估、授权和限制法规：已注册物质信息。https://echa.europa.eu/information-on-chemicals/registered-substances (2025)。访问日期：2025年8月12日。

[25] Wilding, W. V., Rowley, R. L. & Oscarson, J. L. Dippr® 项目 801 评估的过程设计数据。流体相平衡 **150**, 413–420 (1998)。网址 https://doi.org/10.1016/S0378-3812(98)00341-0.[26] 经济合作与发展组织 (OECD)。经合组织现有化学品数据库。https://hpvchemicals.oecd.org/ui/AllChemicals.aspx (2025)。访问日期：2025年10月17日。[27] 美国环境保护署。2020年化学品数据报告 (cdr) 数据 (2020)。网址 https://www.epa.gov/chemical-data-reporting/access-chemical-data-reporting-data。访问日期：2025年12月22日。[28] 北欧部长理事会。SPIN：北欧国家制剂中的物质 (2025)。网址 http://spin2000.net/。访问日期：2025年12月22日。[29] 日本经济产业省 (METI)。日本一般化学物质年度制造和进口数量 (2025)。网址 https://www.meti.go.jp/policy/chemical_management/kasinhou/information/volume_general.html。访问日期：2025年12月22日。

[30] Dysert, L. R. 和 Christesen, P. AACE International 推荐实践第18R-97号：成本估算分类系统——在工艺行业工程、采购与施工中的应用。实践指南，AACE International Inc.，美国纽约（2016年）。网址 https://services.austintexas.gov/edims/document.cfm?id=280770。访问日期：2025年10月17日。

[31] Cullen, L.、Meng, F.、Lupton, R. 和 Cullen, J. M. 减少化学品生产温室气体排放的不确定性。《自然·化学工程》**1**，311–322（2024年）。网址 https://www.nature.com/articles/s44286-024-00047-z。

[32] Parvatker, A. G. 等基于工艺放大和工艺设计计算的二十种麻醉活性药物成分从摇篮到大门的温室气体排放。《*ACS*可持续化学与工程》**7**，6580–6591（2019年）。网址 https://doi.org/10.1021/acssuschemeng.8b05473。

[33] Huber, E., Bach, V., Holzapfel, P., Blizniukova, D. & Finkbeiner, M. 化学品缺失生命周期清单数据的确定方法（RREM）。可持续性**14**, 3161（2022年）。网址 https://doi.org/10.3390/su14063161.[34]Heijungs, R. 概率、统计学与生命周期评价：处理不确定性和敏感性的指南（Springer Nature，瑞士查姆，2024年），第1版。网址 https://doi.org/10.1007/978-3-031-49317-1.[35] Santos, A., Barbosa-Póvoa, A. & Carvalho, A. 化学工业中的生命周期评价——综述。《化学工程最新观点》**26**, 139–147（2019年）。网址 https://doi.org/10.1016/j.coche.2019.09.009。

[36] Althaus, H.-J. 等化学品生命周期清单。最终报告，ecoinvent 数据 v2.0，第 8 号，瑞士生命周期清单中心，瑞士迪本多夫（2007年）。访问日期：2025年10月17日。

[37] Rada, E. C., Schiavon, M. & Torretta, V. 垃圾焚烧发电厂六价铬排放控制的监管策略。《清洁生产杂志》**278**, 123415 (2021). URL https://doi.org/10.1016/j.jclepro.2020.123415.[38]Euro Environmental Ltd. 案例研究：管理废物回收焚烧炉中的六价铬暴露（2024年）。URL https://www.euroenvironmental.co.uk/news/item/managing-chromium-vi-exposure-in-waste-recycling-incinerators。访问日期：2025年10月31日。[39] Astrup, T., Rosenblad, C., Trapp, S. & Christensen, T. H. 废物焚烧空气污染控制残渣中的铬释放。《环境科学与技术》**39**, 3321–3329 (2005). URL https://doi.org/10.1021/es049346q.

[40] 巴塞尔公约秘书处。《陆地焚烧技术准则》(RSM出版社，日内瓦，瑞士，1997年)。URL http://www.basel.int/Portals/4/download.aspx?d=UNEP-CHW-WAST-GUID-IncinerationLand.English.pdf。访问日期：2025年10月24日。

[41] 巴塞尔公约秘书处，联合国环境规划署。《关于控制危险废物越境转移及其处置的巴塞尔公约》：正文及附件（2025年修订版）。联合国环境规划署，巴塞尔公约秘书处（2025年）。URL https://www.basel.int/Portals/4/download.aspx?e=UNEP-CHW-IMPL-CONVTEXT-2025.English.pdf。

[42] Blasenbauer, D. 等。欧洲废弃物焚烧底灰利用的法律现状与当前实践。《废弃物管理》**102**，868–883（2020）。URL https://doi.org/10.1016/j.wasman.2019.11.031。

[43] Wilbur, S.B. 等. 铬毒理学概况（美国卫生与公众服务部、公共卫生服务局，亚特兰大，GA，2012年）。网址 https://www.ncbi.nlm.nih.gov/books/NBK158858/。访问日期：2025 年 11 月 11 日。

[44] 国际协会联盟。《关于消耗臭氧层物质的蒙特利尔议定书》。技术报告，国际协会联盟（1987年）。网址 https://uia.org/s/or/en/1100031640。访问日期：2025年10月24日。

[45] Jacob, P.-M. & Lapkin, A. 有机化学网络的统计学。反应化学与工程**3**，102–118（2018）。URL https://doi.org/10.1039/C7RE00129K。

[46] Greed, S. 将溶剂变成其他东西。《自然评论：化学》**6**，519（2022）。URL https://www.nature.com/articles/s41570-022-00414-5。

[47] Zuiderveen, E. A. R. 等新兴生物基产品减少环境影响的潜力。《自然·通讯》**14**，8521（2023）。网址 https://doi.org/10.1038/s41467-023-43797-9。

[48] Lee, S. Y. 等用于生产生物基化学品的综合代谢图谱。《自然·催化》**2**，18–33（2019年）。网址 https://doi.org/10.1038/s41929-018-0212-4。

[49] Rahimi, A. & García, J. M. 废塑料的化学回收及其在新材料生产中的应用。《自然评论：化学》**1**（2017年）。网址 https://doi.org/10.1038/s41570-017-0046。

[50] Bustillo, L., Laino, T. & Rodrigues, T. 化学领域自动化好奇心驱动发现的兴起。《化学科学》**14**，10378–10384（2023）。网址 https://doi.org/10.1039/D3SC03367H。

[51] Satopaa, V., Albrecht, J., Irwin, D. & Raghavan, B. 大海捞"针"：检测系统行为中的拐点。载于*2011*年第*31*届分布式计算系统研讨会国际会议（*ICD- CSW*），166–171（IEEE，2011年）。URL https://ieeexplore.ieee.org/document/5961514。访问日期：2025年-10- 17。

致谢

本出版物作为NCCR Catalysis（资助编号180544和225147）的一部分完成；NCCR Catalysis是由瑞士国家科学基金会资助的国家研究能力中心。S.C.还感谢USorb-DAC项目提供的资助，该项目由格兰瑟姆环境保护基金会向RMI的气候技术加速器项目Third Derivative提供拨款支持。我们感谢IBM的Teodoro Laino博士及其团队的支持，他们致力于RXN工具的开发。我们感谢Carbon Minds的Arne Kätelhön在稿件审阅期间提出的宝贵意见和建议。M.P.感谢欧盟地平线2020研究与创新计划对CIRCULAR FOAM项目的资助，资助协议编号为101036854。本工作仅反映作者的观点，并不代表欧盟委员会的观点；欧盟委员会不对文中所含信息可能被用于任何用途承担责任。

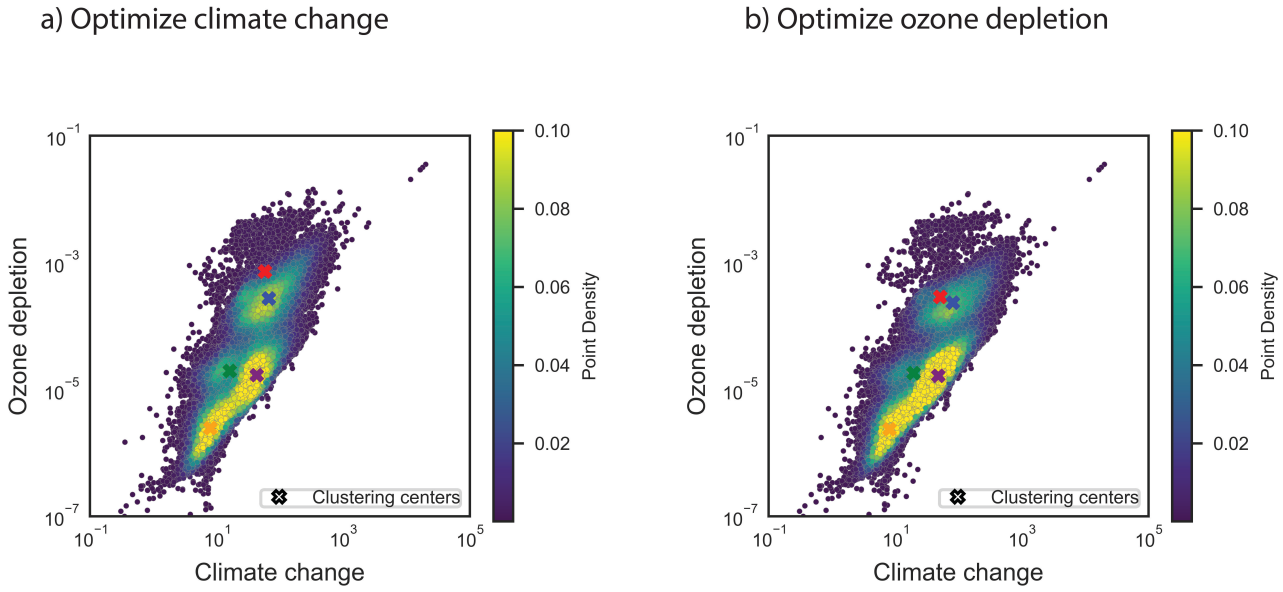
作者贡献

S.C. 参与了研究概念化、方法开发、数据整理和验证，并撰写了手稿初稿。T.L. 协助开发方法论。J.N. 协助进行数据

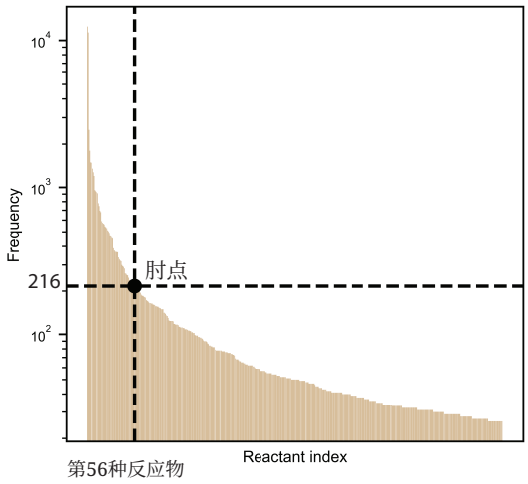
分析并审阅了手稿。C.O. 参与了验证以及数据整理，并协助开发方法论。M.P. 协助进行方法开发，参与用户界面开发，并审阅了手稿。J.S. 参与了研究概念化和数据发布，协助进行方法开发和手稿撰写，并负责项目监督。A.B. 参与了研究概念化和方法开发，协助撰写手稿，负责项目监督，并获得了资金支持。所有作者均参与了讨论并完成了手稿的定稿。

利益冲突

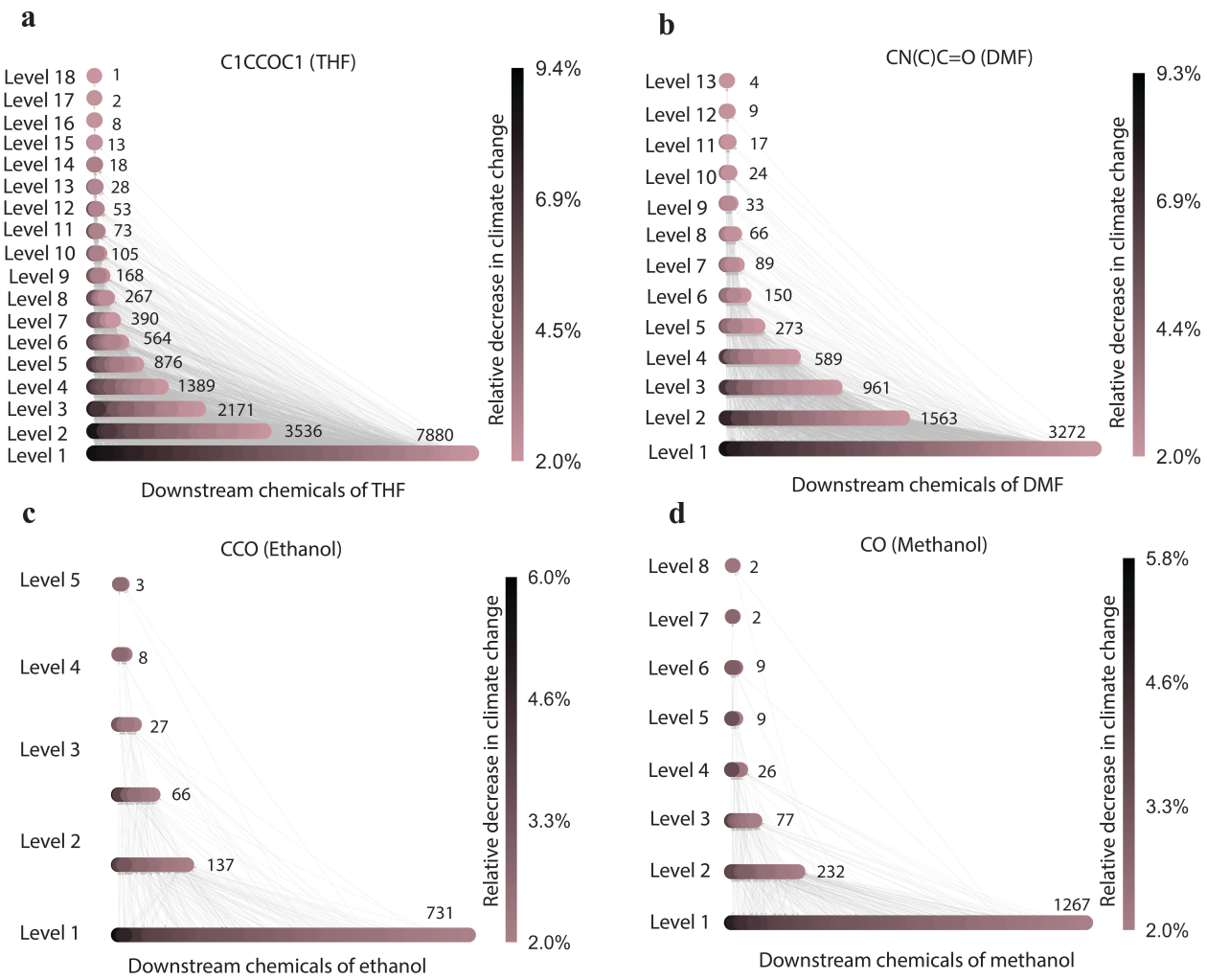
作者声明以下财务利益和个人关系可能被视为潜在的利益冲突：A.B. 曾在埃克森美孚和道达尔能源的研发评审委员会任职，这两家公司均涉足石油和天然气以及化学品生产。A.B. 持有为工业界提供服务的企业的所有权权益，其中一些企业可能在化学工业领域运营。具体而言，A.B. 持有 Carbon Minds 的所有权权益；该公司提供生命周期评价数据库，除其他用途外，也用于本研究中验证 CRYSTAL 模型。T.L. 在完成对本研究的贡献后，加入 Carbon Minds 担任员工。其余作者声明，不存在可能影响本文所报告研究的已知竞争性财务利益或个人关系。



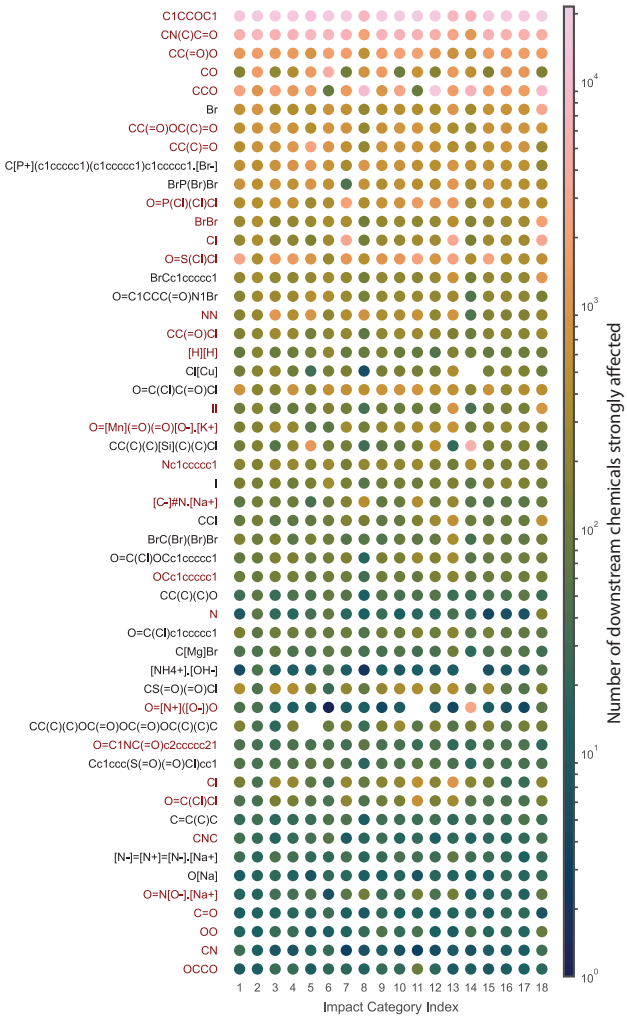
扩展数据图1：a）使用高斯混合聚类算法，沿气候变化优化路径对化学品进行聚类所得的五个簇的聚类质心。**b**）沿臭氧消耗优化路径对化学品进行聚类所得的五个簇的聚类质心——详情见补充文本。



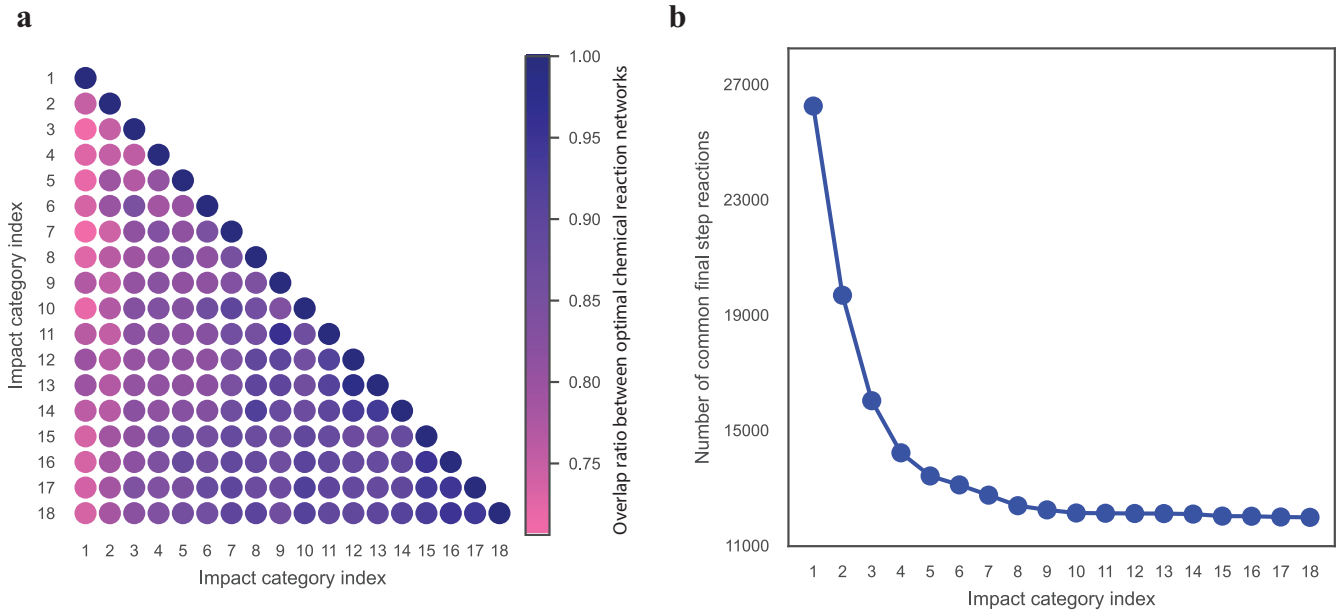
扩展数据图2：最常用作反应物的化学品（$n = 56$），使用肘部法确定[51]。



扩展数据图 3：枢纽化学品对下游产品的气候变化影响。 每种枢纽化学品的气候变化影响降低 10% 后，下游气候变化影响的相对降幅：a）四氢呋喃（THF），b）二甲基甲酰胺（DMF），c）乙醇，d）甲醇。



扩展数据图4：热图显示了枢纽化学品在多个环境影响类别中的敏感性。在各环境影响类别中，每种枢纽化学品减少 10% 时，影响降低超过 2% 的强受影响下游化学品数量。环境影响类别的完整名称如下：1. 酸化：陆地，2. 气候变化，3. 生态毒性：淡水，4. 生态毒性：海洋，5. 生态毒性：陆地，6. 能源资源：不可再生化石能源，7. 富营养化：淡水，8. 富营养化：海洋，9. 人体毒性：致癌，10. 人体毒性：非致癌，11. 电离辐射，12. 土地利用，13. 材料资源：金属/矿物，14. 臭氧消耗，15. 颗粒物形成，16. 光化学氧化剂形成：人体健康，17. 光化学氧化剂形成：陆地生态系统，18. 水资源使用。



扩展数据图5：影响化学反应网络中最优路径选择的环境影响类别。a）针对18 种环境影响类别优化的化学反应网络之间的重叠比率。**b**）随着纳入考虑的环境影响类别数量增加，最优网络之间共有反应的数量减少。对于超过11000 种化学品，化学反应网络优化可确定一个能同时最大限度降低所有环境影响的最终反应步骤；而对于其余化学品，不同影响类别之间仍存在权衡。环境影响类别的全称如下：1. 臭氧消耗，2. 富营养化：海洋，3. 水资源使用，4. 材料资源：金属/矿物，5. 生态毒性：陆地，6. 土地利用，7. 电离辐射，8. 能源资源：不可再生化石能源，9. 酸化：陆地，10. 富营养化：淡水，11. 颗粒物形成，12. 光化学氧化剂形成：人体健康，13. 光化学氧化剂形成：陆地生态系统，14. 气候变化，15. 生态毒性：淡水，16. 生态毒性：海洋，17. 人体毒性：非致癌，18. 人体毒性：致癌。仅纳入网络中至少具有两条反应路径的化学品，以进行严格评估。