

尽管目前模型仅在红队测试中出现勒索、威胁、拒绝被关闭等行为,但当下向“代理型AI”(即AI系统能够在现实世界中自主完成任务)发展的趋势,可能暴露出更多行为失范倾向,并揭示出AI系统易受黑客攻击、网络钓鱼等网络安全威胁的脆弱性。

这些问题的解决方法尚不明朗。一种普适的方法是提升AI素养,并要求使用AI系统的所有人时刻保持警惕,因为他们的请求可能会引发误导、伪造等失范行为,且AI代理可能做出潜在的危险行为。尽管已有大量研究正在攻克如何从技术上解决这些问题,但目前仍缺乏有效的预防方法。Anthropic公司首席执行官Dario Amodei在一篇文章中指出,除非人类能够更好理解这些模型的内在运行机制,否则此类问题将无解,但令人唏嘘的是,哪怕是负责设计和训练模型的工程师也对此知之甚少。

包括我自己在内的一些研究者都认为,这些问题带来的风险已经不容忽视,一篇论文就此指出,“完全自主化的AI代理不应被开发”。换言之,AI系统必须始终处于人类的控制和监督下。然而,这种限制很可能违背多数AI公司的商业利益,也不符合美国当前的政治环境——在提升国家工业竞争力的目标面前,任何政府监管都显得无足轻重。如前文所述,虚构角色“高级人工智能Alex”过分执着于这一目标,可能会加剧实际开发的AI与最符合社会利益的AI之间产生的行为偏差。

(译自《Science》,2025,389(6758))

改造沼泽可能有助于遏制甲烷排放,但负面影响尚不明确

随着强效温室气体浓度持续攀升,科学家正“谨慎”尝试新的地球工程手段

文/Paul Voosen

湿地因其水流滞缓、水体缺氧且富含有机淤泥,成为厌氧微生物的理想栖息地,而厌氧微生物会释放甲烷这一强效温室气体。气候变暖似乎正在加速这种微生物活动,促使全球甲烷浓度激增,进而加剧了气候变暖,形成一种危险的反馈循环。为打破这一循环,一些研究者正在探索一种非常规设想:改造沼泽。

2025年启动一项“甲烷排放反馈效应研究与行动”(Feedback Research and Action on Methane Emissions, FRAMES)项目,拟通过化学手段改造湿地微生物生态系统,以遏制甲烷排放。该项目计划从实验室开始,逐步扩展至小规模实地试验。研究人员指出,这种新的地球工程手段可能带来不可预见的后果,探明这些后果是开展该项目的主要目的。美国环保协会(为该项目提供支持的非营利组织)气候创新倡议联合负责人Brian Buma表示,“这需要谨慎行事。我们正身处未知领域。”

过去20年,大气甲烷浓度已上升近10%,据称在引起全球变暖的气体中,甲烷占1/3。研究人员了解到,甲烷浓度上升的主要来源并非石油和天然气行业,因为大气甲烷中碳12(生物活动偏

爱的轻同位素)的丰度持续增加。

畜牧业扩张和垃圾填埋场增加固然对此有所影响,但研究人员发现湿地的甲烷排放量也在上升。全球变暖正在改变降水模式,导致某些地区沼泽面积扩大,而永久冻土层融化形成了新的湿地。气候变暖也为产甲烷菌提供了有利生长环境:气温升高加速其新陈代谢,而水温升高导致溶解氧下降,使这些厌氧生物更容易在生态系统中占主导地位。

2024年,美国能源部的一个科学小组发现,在2002年到2021年,欧亚大陆北部和北美地区湿地的甲烷排放量增加了9%。在2025年5月发表于《Nature》的一项研究中,研究者详细探究了长期监测点的甲烷浓度季节性起伏变化,发现全球湿地甲烷排放量自20世纪80年代以来持续上升。2项研究均将这一趋势与全球变暖相关联。爱丁堡大学大气化学家Paul Palmer表示,“这一切都在无可避免地发生。”

一般而言,遏制甲烷排放的最佳方式首先是阻止湿地持续变暖。科尔盖特大学生态学家Emily Ury表示,这种方式要求减少使用化石燃料。2024年,她与Buma共同发表了一篇综述,阐述了缓解湿地反馈效应的可能方

案。她指出,尽管尚无法落实具体行动,但至少应对人工干预措施进行研究。

遏制湿地甲烷排放有一个著名的方法,这种方法源自早先经历的环境危机。20世纪,燃煤电厂向大气排放大量硫化物,导致形成酸雨。这些排放物对人类和生态系统健康造成了损害,但同时,酸雨的产生也使沼泽的甲烷排放量显著降低。研究人员发现,这是由于系统中过量的硫酸盐使消耗硫酸盐的细菌在竞争中胜过产甲烷菌,导致后者因营养匮乏而减少。

如今,这一经验在东亚地区的稻田种植中得到印证:农民为缓解作物盐胁迫添加了富含硫酸盐的石膏,却偶然间减少了甲烷排放。其他添加剂,如硝酸盐、铁和生物炭(一种用作土壤改良剂的木炭)等同样具有应用潜力。

为推进 FRAMES 项目进程, Ury 与康奈尔大学生态学家 Meredith Holgerson 正在建立一组土壤核心样本(湿地培养罐),以测试这些方案的可行性。Ury 指出,湿地具有高度的异质性,即使某种处理方法对某类湿地有效,但对其他类型可能完全无效,“我们需要以一种真正可复现的方式开展这些实验。”

硫酸盐实地试验最早可能于

2025年同步启动。多伦多大学水文学家 Irena Creed 表示,其合作伙伴之一环保组织 Ducks Unlimited 正迫切希望验证该方法能否遏制甲烷排放,以便应用于其在加拿大管理的已修复湿地。这类试验可能涉及对沼泽化学特性、微生物群落及温室气体排放的基线研究,还需要对使用硫酸盐处理后的湿地进行持续监测。Creed 的研究范围还包括加拿大西部草原坑塘区的城市和农田湿地,并计划在未来 18 个月内启动相关试验。

Creed 指出,在进行这些化学改造时,重点需监测其对生态系统有哪些潜在的负面影响。例如,有研究表明,硫酸盐会加速微生物将污染物汞转化为甲基汞的速率,而甲基汞是一种可在水生物种中生物积累的毒素。对此, Creed 强调,“连续施用 12 h 添加剂并不能一劳永逸地解决问题。”

来自 Spark Climate Solutions (一家专注甲烷减排的小型非营利组织)的生物地球化学家 Dani-

elle Potocek 指出,减少沼泽甲烷排放还有许多可行的方法,例如,可以通过挖掘渠道将沿海沼泽与富含硫酸盐的海洋重新连接,或者可以提高沼泽水位,从而阻止甲烷逸入大气。另一个关键因素是植被类型:空心的莎草和禾草会像“吸管”一样加速甲烷释放。此外,研究表明水生蕨类植物能够促进湿地土壤的氧化过程,进而减少产甲烷菌的数量。

Potocek 补充道,同许多地球工程手段一样,“规模化是关键问题。”定期向湿地投放大量化学药剂不仅成本高昂,且操作难度很高,而非洲和南美洲拥有广袤的天然湿地,这里是甲烷反馈效应的主要发生地,但没有人愿意干扰破坏。

Buma 认为,或许可以小规模实践为起点,将这些经验应用到湿地修复这一全球普遍发展的领域。他表示,“即使只在少数项目中落实甲烷减排,也能为缓解气候变暖带来巨大的附加效益。”

(译自《Science》,2025,389(6758))



苏丹南部苏德地区广袤的热带沼泽地排放的气体,促使全球甲烷水平升至新高

(图片来源:《Science》)