

网站地图 (<http://www.imech.cas.cn/serv/wzdt/>) |

联系我们 (http://www.imech.cas.cn/serv/lxfs/201212/t20121205_3698646.html) |

所刊 (<http://www.imech.cas.cn/serv/szxx/>) | 所内网 (<https://ioa.imech.ac.cn>) |



<https://mail.imech.cas.cn/> | [English \(<http://english.imech.cas.cn/>\)](http://english.imech.cas.cn/)

Institute of Mechanics, Chinese Academy of Sciences

中国科学院 (<http://www.cas.cn/>)

(<http://www.imech.cas.cn/>)



当前位置：首页 (../..../..../) >> 科学传播 (../..../..../) >> 力学园地 (../..../) >> 释疑解惑 (../..../)

【释疑解惑】从轨道杂化角度看干酪根熟化过程——兼谈机器学习的应用

2023-06-02 15:33

【放大 缩小】

《力学园地》编辑部：

你们曾介绍过力学研究所非线性力学国家重点实验室的赵亚溥研究团队在页岩油开发利用领域的若干研究成果，特别是在干酪根分子结构及其熟化过程的探索方面。还能介绍他们在相关领域的最新进展吗？

一名在读的研究生

2023年2月9日

从杂化轨道角度看干酪根熟化过程——兼谈机器学习的应用

马骏

我国原油供应的对外依存度高，是被人卡住的最细的脖子。我国页岩油气潜力巨大，是重要的战略能源，但限于技术与生产水平，对其开发利用率不高。因此需要在关键技术上全力攻坚，实现页岩油革命。干酪根是页岩油气的生成母质和主要赋存介质，是一种不溶于有机溶剂的大分子有机质，具有结构复杂、非均质等特点。可以想象吗？一些固体粉末经过热解之后能够分解成油气混合物，而这些油气就是一种能够改变世界格局的非常规能源！但是，通常干酪根分子量达到了几百、上千，结构难以被准确描述与重构，因此重构干酪根分子结构一直

是科研工作中的重点科学难题。力学所非线性力学国家重点实验室赵亚溥研究团队立足于力学，运用力-化-热-电-光等耦合物理场的理论以及实验、数值模拟、人工智能等方法，成功地重构了干酪根大分子结构。图1中展示了他们构建的16种不同的干酪根大分子。

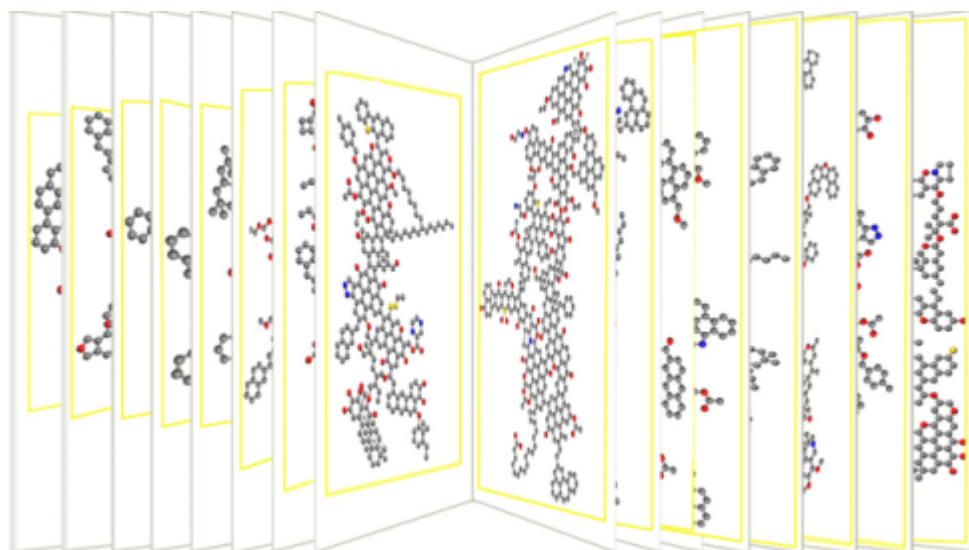


图1 赵亚溥团队构建的干酪根分子群

干酪根是由英文“kerogen”音译而来的，最初的意思是指页岩中能够生成油气的物质。它是动植物残骸经过亿万年的演化形成的，根据演化程度，科学家定义了一个参数“成熟度”，成熟度越高，干酪根越“成熟”，演化程度越高，距离最终产物越近。所以，成熟度是评价页岩油气生成潜力的重要指标，是研究干酪根熟化过程的必要参数。干酪根熟化过程其实就是成熟度不断升高的过程，所经历的地质生成阶段有成岩、深成作用和准变质作用等三个阶段。成岩阶段很好理解，就是早期动植物残骸在外部环境作用下埋入地层形成岩石的阶段，有机质（蛋白质、碳水化合物）在此期间被微生物活动破坏，生成的组分逐渐聚合形成干酪根。而深成作用阶段，则是这些动植物残骸成岩之后，由于埋藏深度的进一步增加，在高温高压的环境中不断“成熟”产生了油气的阶段。在最后的准变质作用阶段，干酪根的成熟度已经很高了，残余干酪根组分转化为石墨碳。从表面上看，成熟度与干酪根中的H/C和O/C原子比直接相关，实际上成熟度演化的本质是干酪根分子结构在变化，也就是分子中原子间的化学键在变化。具体而言，由于温度升高，不稳定的化学键（如单键、双键等）首先断裂，相对稳定的苯环键被保留。这样，干酪根就从最初的大分子热解成油气小分子，如图2所示。

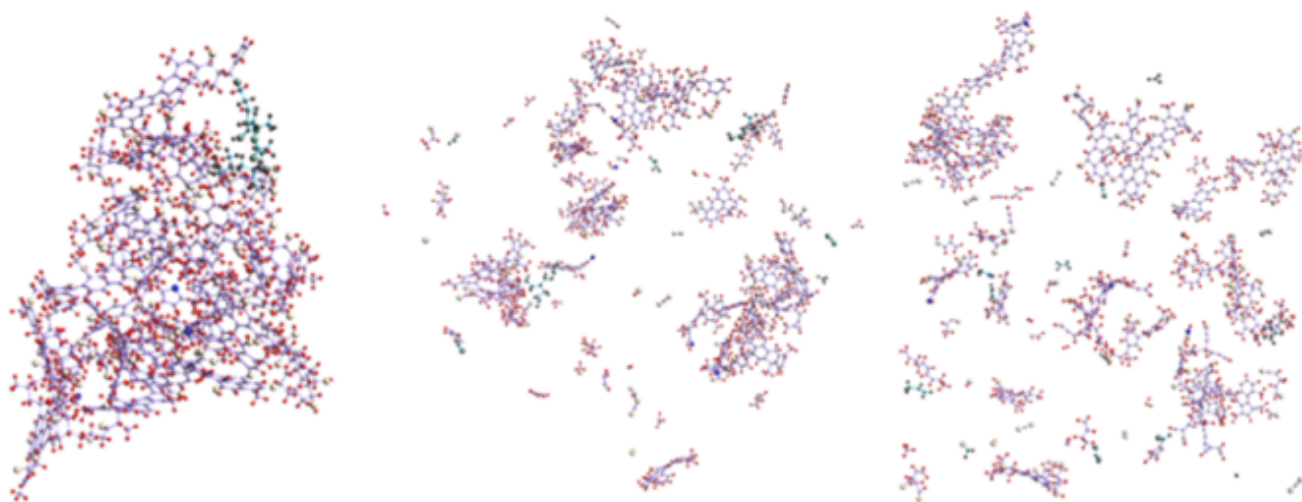


图2 干酪根分子热解过程

这里就要更具体地了解一下化学键是怎么回事儿啦！化学键的概念是科学家在总结长期实践经验的基础上建立和发展起来的，用来概括观察到的大量化学事实，特别是用来说明原子是如何结合成具有确定几何形状的、相对稳定和独立的、性质与其组成原子完全不同的分子。在分子结构的二维平面图中，化学键由相互结合的两个原子之间的短线作为化学键的符号，如图3所示：一根线表示单键，两根线表示双键，正六边形的边表示苯环键。中学的化学课程里给出了如下的定义：化学键是使物质分子内相邻两个或多个原子（或离子）相结合的作用力，它有离子键、共价键和金属键三种类型。我们研究的干酪根分子的化学键属于共价键类型。20世纪初，美国科学家G. N. 路易斯建立了化学键的电子理论。这个理论认为，原子核外的电子沿着一定的轨道运动（严格地说，它表示电子云的形状），原子在形成分子时，外层电子发生了重新分布（转移、共用、偏移等），从而产生了正、负电性间的强烈作用力。这样就揭示了化学键的本质。基于这个理论，共价键是通过原子间共用一对或多对电子（电子云重叠）所形成的。在化合物分子中，两个原子间以共用一对电子而构成的共价键，就是单键；两个原子共用两对电子对，就是双键；两个原子间由三对共享电子组成的重键，便是三键。但是，路易斯的理论无法解释分子的三维构型以及电子轨道形状、分布等等，因此有了轨道杂化理论。

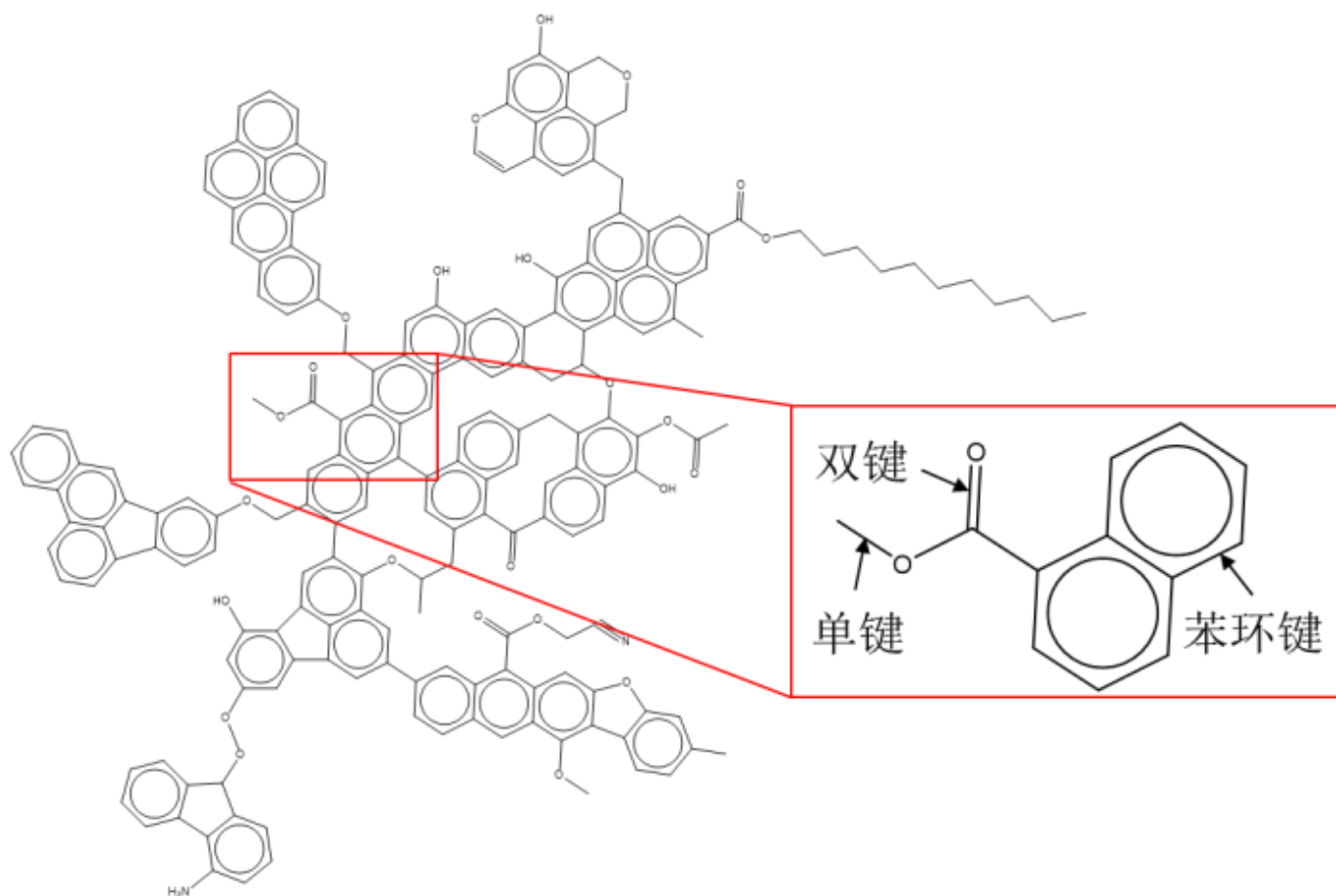


图3 干酪根分子中的单键、双键、苯环键示意

20世纪20至30年代，美国的L. C. 鲍林等人在量子力学基础上提出了轨道杂化概念，使得化学键的理论解释日趋完善。量子理论认为，电子具有量化的能级，不同能级的电子占据不同的原子轨道。它是这样分类的：（1）按照电子与原子核的距离，分为不同的层级（以数字1, 2, 3, 标记）。第1层级距原子核最近，能级最低；

距离越远（层级数越大），能级越高；（2）按照原子轨道的形状，分为 s 轨道（圆形）、p 轨道（哑铃形）、d 轨道（花瓣形）等等，相应的能级逐次提高。轨道杂化理论认为：在成键过程中，由于原子间的相互影响，相邻原子中几个能量相近的不同轨道（更严格地讲是原子轨道的波函数，但它涉及深奥的量子力学，这里就不展开说明了）可以进行线性组合，重新分配能量和确定空间方向，组成数目相等的新轨道。这种轨道重新组合的方式称为“杂化”，杂化后形成的新轨道称为杂化轨道，杂化轨道比原来轨道的成键能力强，所形成的键更加稳定。轨道杂化一般发生在 s 轨道与 p 轨道之间，简称为 sp 型杂化。这类杂化轨道种类很多，最常见的是 sp^1 ， sp^2 和 sp^3 杂化，分别对应的是化学键三键、双键/苯环键和单键。因此，运用轨道杂化方式去解释干酪根熟化过程有助于我们认识干酪根熟化过程的内在原因——分子结构改变。我们下面就来探寻轨道杂化理论在干酪根成熟度问题里的应用吧！

力学所非线性力学国家重点实验室赵亚溥研究团队此前提出的干酪根成熟度评价指数：分子成熟度指数 (Molecule-Maturity Index, MMI) 已被证明能够反映干酪根的生烃潜力，并被许多学者引用表示成熟度。大家知道，烃是碳氢化合物的总称。这里所说的“生烃潜力”便是从理论上给出干酪根最多能够热解产生的碳氢化合物的产量。因此，干酪根的生烃潜力越大，说明越有开采的价值。在此基础上，为了能从分子结构的角度去解释干酪根的熟化过程，赵亚溥团队进一步研究以轨道杂化参数来表征干酪根成熟度，这里的轨道杂化参数指的是分子中具体发生 sp^2 杂化和 sp^3 杂化的碳原子数与氧原子数。比如，苯环中的碳原子是 sp^2 杂化，单键中的碳原子是 sp^3 杂化，统计一个有机大分子中有多少个这样的杂化原子，求和得到杂化总数，作为衡量轨道杂化的参数。当一个原子从 sp^2 杂化转为 sp^3 杂化时，化学键通常是由双键转为单键，而且温度越高，结构越不稳定。按照这个规律，人们也认为干酪根在热解过程随着温度的升高，双键转为单键， sp^2 杂化转为 sp^3 杂化。然而事实真是如此吗？不少学者曾注意到高成熟度的干酪根分子结构会趋于芳香结构的现象，赵亚溥团队则依据杂化理论给出了科学的解释。其实干酪根中的 sp^2 杂化的碳原子主要是苯环上的碳原子，由于苯环键是一种十分稳定的化学键，所以在热解时并不容易断裂，反而导致热演化过程中 sp^2 杂化比例不降反升。事实上，油气的生成乃是来源于脂肪族支链也就是非苯环部分的断裂，这部分的化学键通常是单键，也就是 sp^3 杂化碳，断裂之后生成了油气小分子。随着温度的升高，芳香结构（也就是苯环结构）被保留了下来，于是干酪根趋于“成熟”。这里，简单说明一下，什么是脂肪结构？什么是芳香结构？一般而言，一个有机大分子，它的主体结构要么是长链，要么是环。长链结构的主体单元通常是由碳碳单键组成，这样的结构我们称为脂肪结构，这里的碳被我们称为脂肪碳。而以正六边形的苯环为主体的结构，通常由多个环相连，这样的结构称为芳香结构，这里的碳被称为芳香碳，这是因为曾经发现的一种这类化合物具有芳香气味。

那么，这些分子结构信息是如何得到呢？可以采用不同的实验方法获得分子结构信息，如元素分析、X 射线光电子能谱、核磁共振 (NMR) 等技术，其中 NMR 方法可以获得分子结构中的官能团信息，被广泛用于确定干酪根分子结构。然而，要想准确地分析 NMR 谱图对应的分子结构信息，需要专业的知识与人员，这对只需获取个别的干酪根分子结构来说没有什么问题，但若想大量地构建干酪根分子模型，就需要耗费很多的人力物力。机器学习，因它在处理大数据方面有独特的优势，能够解决许多复杂的问题，近年来常常被研究人员采用。赵亚溥团队创新性地采用机器学习方法对干酪根分子进行重构，通过大量数据的训练“学习”特征，进而对未知分子进行重构，平均相似度达到了90%以上。这说明了机器学习智能化构建干酪根分子的优越性，提高了科学研究的效率。那么，机器到底是如何预测出新的分子结构的呢？具体而言，他们首先将已知的分子结构和相应的NMR谱图输入机器学习模型，这时要为每个谱图数据标注其表示的分子结构（换言之，做好“标签”）；其次，对机器进行训练。通过机器自身的学习功能，经过训练的机器就学习到了谱图各个细节和分子结构特征的关联，并把它们

存为资料数据库；然后，输入新获取的未知分子的谱图，机器便利用其资料库中的数据，预测出所对应的新分子结构。当然，完成这类任务，功能强大的计算机的效率要比人工甄别与试错的效率高了许许多多，一般只需要几十分钟到几小时便可给出预测的结构，而且还可以同时对多组数据进行处理。图4给出机器学习预测性能随训练迭代次数变化情况。可以看到：当迭代次数为136次时，机器预测的分子结构与目标分子便达到了完全的相似。

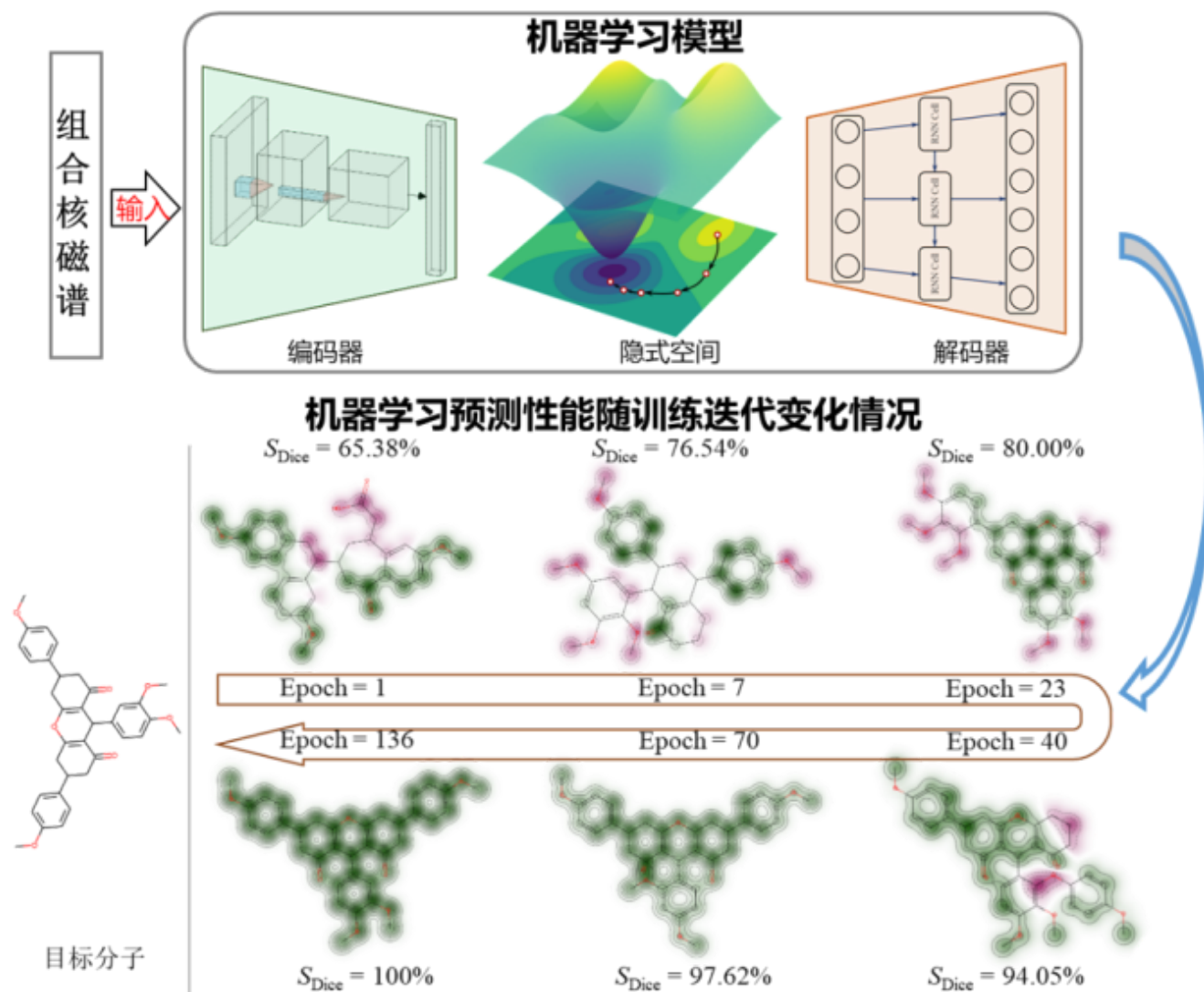


图4 基于机器学习的干酪根分子重构

机器学习除了获取干酪根的分子结构以外，还可以用于描述干酪根的成熟度。前面提到，赵亚溥团队曾提出一个“分子成熟度指数MMI”，它是基于干酪根分子式中的元素氢（H）、氧（O）、氮（N）和硫（S）与碳（C）之比（H/C，O/C，N/C，S/C）来表示的。随着研究的深入，他们又提出基于轨道杂化的干酪根成熟度指标“轨道杂化成熟度指标”（Orbital Hybridization Maturity Index, OrbHMI）。指标OrbHMI是以轨道杂化信息为衡量参数的，这个轨道杂化参数既可以通过分析分子结构得到，也可以利用机器学习而获得。它能够直接反映干酪根分子结构变化的深层次原因，也就是具体到化学键的断裂与重组。图5给出了基于机器学习和量子力学的干酪根成熟度表征模型。它利用 NMR 谱图提取原子的轨道杂化方式并统计 sp^2 ， sp^3 杂化原子数，这里之所以不统计 sp^1 杂化原子数，是因为三键太不稳定，已知的干酪根分子中三键的比例非常低，缺少统计价值。

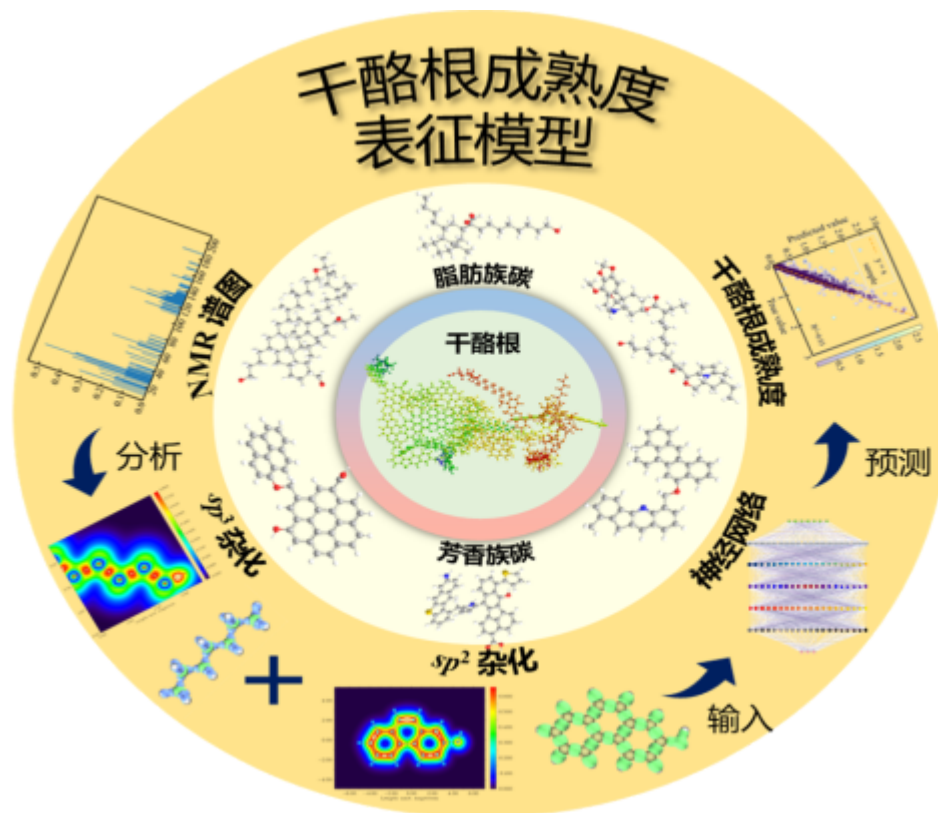


图5 基于量子力学和机器学习的干酪根成熟度表征模型

赵亚溥团队采用量子力学的轨道杂化概念，阐明了干酪根熟化过程是 sp^2 杂化碳占比升高、 sp^3 杂化碳占比降低的过程，因而干酪根分子结构由脂肪结构逐步向芳香结构转变。他们还进一步揭示了机器学习模型的预测结果平均误差低于5%，超过87%的预测误差在10%以内，从而确认了机器学习对干酪根成熟度预测的高精确度。图6和7将 OrbHMI 与 MMI 以及机器学习预测结果进行了对比，发现彼此间吻合良好，表明了轨道杂化成熟度指标的可信性，也进一步说明了机器学习在表征干酪根成熟度方面的杰出优势。图6比较了两个成熟度指标的线性关系，二者的数值均是根据公式计算得出的，样本点的颜色越深表示 MMI 与 OrbHMI 越相近。类似地，图7中样本点的颜色越深表示机器预测值与OrbHMI 之间的误差越小。

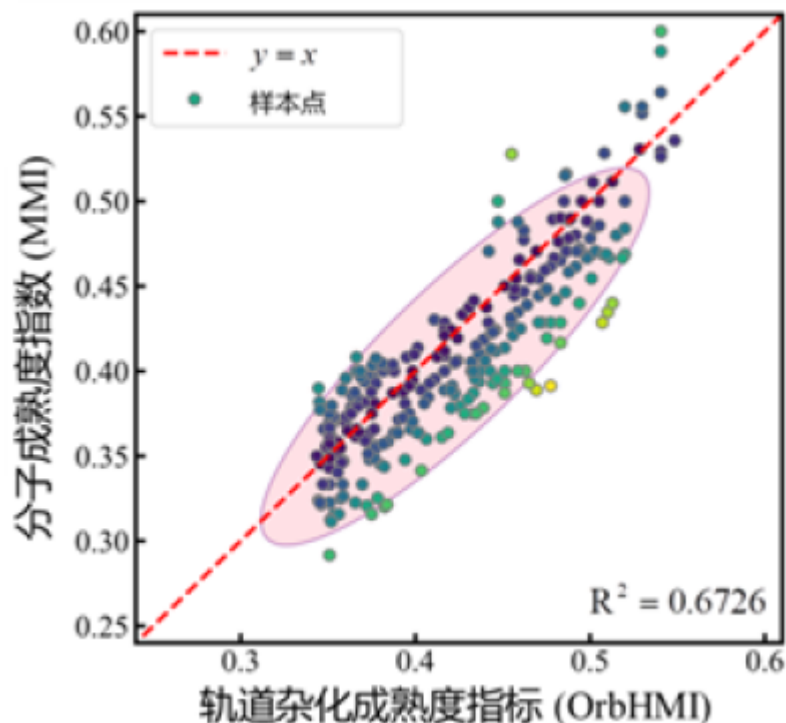


图6 两种成熟度OrbHMI与MMI的比较

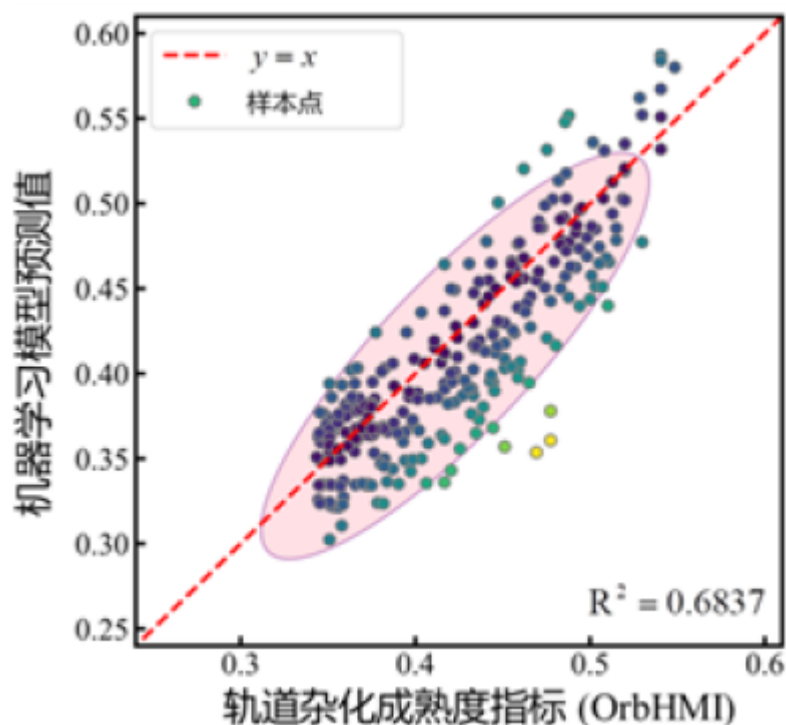


图7 机器学习模型预测结果

赵亚溥研究团队关于机器学习在干酪根模型研究方面的工作引起国内外专家的高度关注。相关成果发表于国际权威学术刊物《Global Challenges》、《Journal of Petroleum Science and Engineering》、《Energy》、《Fuel》、《Energy & Fuels》等。这里仅涉及轨道杂化成熟度指标的研究部分，有兴趣的读者可

通过参考文献获得更多的细节。

参考文献

- [1] Wang X, Huang X, Gao M, Zhao Y P. Mechanical response of kerogen at high strain rates. *International Journal of Impact Engineering*, 2021, 155: 103905. <https://doi.org/10.1016/j.ijimpeng.2021.103905> (<https://doi.org/10.1016/j.ijimpeng.2021.103905>)
- [2] Zhao Y P. Physical mechanics investigation into carbon utilization and storage with enhancing shale oil and gas recovery. *Science China Technological Science*, 2022, 65: 490-492. <https://doi.org/10.1007/s11431-021-1985-y> (<https://doi.org/10.1007/s11431-021-1985-y>)
- [3] Ma J, Kang D L, Wang X H, Zhao Y P. Defining kerogen maturity from orbital hybridization by machine learning. *Fuel*, 2022, 310: 122250. <https://doi.org/10.1016/j.fuel.2021.122250> (<https://doi.org/10.1016/j.fuel.2021.122250>)
- [4] Kang D L, Zhao Y P. Predicting the molecular models, types, and maturity of kerogen in shale using machine learning and multi-NMR spectra. *Energy & Fuels* 2022, 36: 5749-5761. <https://doi.org/10.1021/acs.energyfuels.2c00738> (<https://doi.org/10.1021/acs.energyfuels.2c00738>)
- [5] Wang XH and Zhao Y.-P. The time-temperature-maturity relationship: a chemical kinetic model of kerogen evolution based on a developed molecule-maturity index. *Fuel*. 2020, 278: 118264. <https://doi.org/10.1016/j.fuel.2020.118264> (<https://doi.org/10.1016/j.fuel.2020.118264>)



中国科学院 (<http://www.cas.cn>)
CHINESE ACADEMY OF SCIENCES

中国科学院力学研究所 版权所有 京ICP备05002803号 京公网安备110402500049

地址：北京市北四环西路15号 邮编：100190

(<http://bszs.conac.cn/sitename?method=show&id=081D2D6355AD574EE053022819ACCBA7>)

