



古气候数据同化: “4.85亿年地球表面温度”一文的评论

张仲石^{*}, 刘永岗

北京大学物理学院大气与海洋科学系, 北京 100101

^{*} 通讯作者, E-mail: zhongshi.zhang@pku.edu.cn

收稿日期: 2025-01-05; 收修改稿日期: 2025-03-28; 接受日期: 2025-04-02; 网络版发表日期: 2025-04-14

国家自然科学基金项目(42125502)和国家重点研发计划项目(2024YFF0807903)资助

数据同化(Data Assimilation, DA)是一种融合多源观测数据与数值模型以提高模型模拟质量和优化数据的技术, 其核心思想是在气候模拟的基础上, 利用实际观测数据修正模拟结果的偏差, 同时对观测数据进行自动择优. 简单来说是利用某些数学方法迫使模拟结果更加接近观测数据.

数据同化技术的兴起伴随着数值天气预报的发展. 20世纪50年代, 气象学家认识到数值模型的预测会受到初始条件的影响, 也就是初值问题. 为了攻克这个问题, 气象学家从60年代开始发展数据同化方法优化初始条件. 首先是基于插值的简单方法, 例如最优插值法; 随后是卡尔曼滤波法、变分法、集合卡尔曼滤波法、粒子滤波法等相对复杂的方法. 目前, 集合卡尔曼滤波和粒子滤波是常用的两种方法; 前者适用于高维、非线性和数据丰富的同化, 计算效率较高, 但在处理强非高斯系统时仍存在一定局限; 后者适用于强非线性、非高斯系统, 但在高维问题上计算成本较高, 并面临粒子退化问题. 数据同化也不再局限于传统的气象与海洋领域, 逐渐扩展到其他复杂的环境和生态系统建模中, 例如气候变化预测、污染物扩散模型、地下水流动模型等.

随着古气候数值模拟结果的不断积累, 古气候数

据同化应运而生. 它首先在过去千年的模拟结果中应用(例如, von Storch等, 2000; Goosse等, 2010; Tardif等, 2019; Valler等, 2022), 随后全新世(Erb等, 2022)、末次冰期盛冰期(Tierney等, 2020)、始新世(Tierney等, 2022)等, 时间尺度上逐步从亚轨道向轨道和构造时间尺度拓展. 中国学者在古气候数据同化领域取得重要进展, 近年来已系统开展了过去千年同化模拟(Shi等, 2019; Fang等, 2022; Lyu等, 2024), 并逐步拓展至构造时间尺度同化模拟(Li等, 2024). 最近, 《中国科学: 地球科学》上发表的古气候数据同化相关文章介绍了古气候数据同化的原理与进展(张昊勋等, 2025; 方苗和李新, 2016).

古气候数据同化将是古气候模拟领域未来快速发展的方向. 为了让地球科学领域的学者能够更好地了解古气候数据同化结果的优缺点, 我们评论*Science*期刊上发表的一篇研究论文“A 485-million-year history of Earth’s surface temperature”(4.85亿年地球表面温度)(Judd等, 2024; 后面以J24代替该论文), 并指出论文同化结果可能存在的不确定性.

J24是一篇深时古气候数据同化代表性文章. 这一研究选择显生宙过去4.85亿年中85个时间片段, 针对每个片段开展八次模拟(集合模拟); 然后采用离线集

中文引用格式: 张仲石, 刘永岗. 2025. 古气候数据同化: “4.85亿年地球表面温度”一文的评论. 中国科学: 地球科学, 55(5): 1728–1731, doi: [10.1360/SSTe-2025-0004](https://doi.org/10.1360/SSTe-2025-0004)

英文引用格式: Zhang Z, Liu Y. 2025. Paleoclimate data assimilation: Comments on “A 485-million-year history of Earth’s surface temperature”. *Science China Earth Sciences*, 68(5): 1682–1685, <https://doi.org/10.1007/s11430-025-1558-8>

合卡尔曼滤波方法针对每个片段进行数据同化, 简单来说是利用地质记录对每个片段的集合模拟进行订正并给出置信区间, 获得了过去4.85亿年全球平均表面温度的变化. 研究采用的地质记录来自显生宙海表面温度(PhanSST)数据库收集的溫度重建(Judd等, 2022); 重建方法包括浮游有孔虫、腕足类和软体动物碳酸盐的氧同位素($^{18}\text{O}_{\text{carbonate}}$)、牙形石的磷酸盐氧同位素($^{18}\text{O}_{\text{phosphate}}$)、浮游有孔虫的Mg/Ca、海洋生物脂质(TEX_{86})以及烯酮不饱和指数(U_{37}^K).

文章的研究结果显示, 在过去4.85亿年里, 地球表面平均温度在11~36℃波动, 整体呈现出三个主要的长期趋势: 从早奥陶纪到晚石炭纪(约4.85亿年前到3亿年前)温度整体呈下降趋势; 从晚石炭纪到晚白垩纪(约3亿年前到9000万年前)温度整体呈上升趋势; 从白垩纪晚期到现在温度又呈现下降趋势. 同化结果可以很好地再现一些重要的事件, 例如奥陶纪-志留纪之交(约4.45亿年前)的显著降温和二叠纪-三叠纪之交(约2.51亿年前)的急剧升温. 研究进一步将显生宙地球气候划分为寒室(Coldhouse)、冷室(Coolhouse)、过渡(Transitional)、暖室(Warmhouse)和热室(Hothouse). 显生宙以过渡和暖室气候为主(占比50%), 目前我们所处的寒室气候在显生宙并不常见. 研究进一步考虑纬向温度梯度、热带温度以及气候敏感性的演化. 极地放大效应往往在暖期大幅度减弱极地赤道之间的经向温度梯度. 热带表面温度可以在22~42℃波动, 因此并不支持热带温度恒定的观点. 显生宙地球系统敏感性达到8℃, 远高于现代气候的2~5℃. 这一研究为我们理解地球长期气候变化提供了框架.

然而, 在评价同化结果时, 必须注意到, 由于地质证据极强的时空不均匀性, 古气候数据同化具有不确定性. 数据同化, 例如集合卡尔曼滤波, 需要设定每个记录点的影响范围(同化半径). 换句话说, 每个记录点能够改变其周围多少个格点数据, 这是提前设定的. 在这一范围内, 记录点对模型格点数据的修正能力也会随距离减弱, 这一减弱的幅度也是提前设定的.

J24一文的作者特别指出, 以下6个时段因数据量过少, 他们进行了前后两个时期地质数据的融合, 分别为: (1) 古新世塞兰特期(Selandian, 61.66~59.24Ma)和白垩纪丹尼期(Danian, 66~61.66Ma); (2) 侏罗纪辛涅缪尔期(Sinemurian, 199.46~192.9Ma)和赫塘期(Het-

tangian, 201.36~199.46Ma); (3) 三叠纪安尼期(Anisian, 246.7~241.46Ma)和奥伦尼克期(Olenekian, 249.88~246.7Ma); (4) 二叠纪卡匹敦期(Capitanian, 264.34~259.55Ma)和沃德期(Wordian, 269.21~264.34Ma); (5) 泥盆纪布拉格期(Pragian, 412.4~410.5Ma)和洛赫考夫期(Lochkovian, 419~412.4Ma); (6) 志留纪埃隆期(Aeronian, 440.49~438.59Ma)和鲁丹期(Rhuddanian, 443.07~440.49Ma). 在这一工作中, 可以用于同化的数据数量在时间上其实是不均匀的, 从埃隆期和鲁丹期的5个到全新世的150个.

一旦可同化的重建数据过少, 加之重建数据的空间分布不均, 古气候数据同化结果可能存在显著的模型依赖性. 例如埃隆期和鲁丹期, 因可同化的重建数据只有5个, 计算的全球平均温度几乎完全由模拟结果来决定(图1). 整个PhanSST数据库中的重建海表面温度空间分布显著不均, 数据点主要集中在现今北美、欧洲和澳大利亚等地区, 而在南半球和赤道地区相对较少. 另外, 大部分气候模型模拟的海表面温度都有顽固的系统偏差(Zhang等, 2023), 尤其是西风带和大陆西海岸附近; 如果重建记录处于这些区域(如图1), 重建很可能只反映局地状况, 其约束反而可能导致其他区域温度的过度校正, 增大同化结果的误差. 因此, 重建数据的空间分布不均叠加气候模型的系统偏差可能会使同化结果产生相当大的误差.

此外, 可同化的重建数据如果时间分辨率很低, 轨道尺度气候变率难以分辨, 有可能使同化结果显著偏离该时段的平均气候. 重建的数据可能是完全不同地球轨道参数条件下的结果, 将这些重建同化在同一个时段中, 会对气候模拟产生错误的约束. 在不同的地球轨道参数条件下, 中高纬年平均海温的变幅可以超过2℃, 在寒室气候中可以超过5℃. 当然, 重建数据在时间上插值可以在一定程度上解决这一问题.

需要说明的是, 离线古气候数据同化通常订正地质记录反映的气候变量(例如气温、降水量、海表面温度、海表面盐度), 并往往假设地质记录与这些气候变量之间存在线性关系; 同时会利用协方差等数学方法订正大气、海洋环流的动力场, 但这种订正因缺少动力学约束难免引入新的误差. 因此, 利用离线古气候数据同化结果理解过去气候变化的机理需要审慎地考虑其不确定性. 当然, 离线古气候数据同化的某些不足可以用在线古气候数据同化(例如在线集合卡尔曼

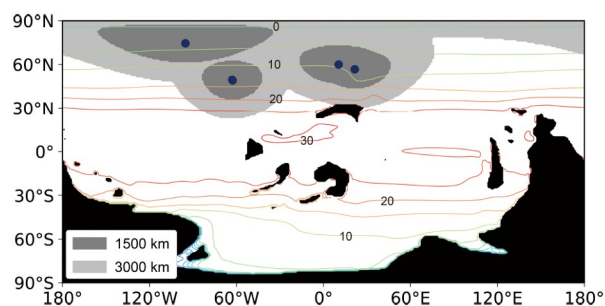


图1 鲁丹期记录点的同化范围

深灰色和灰色分别表示同化半径1500和3000km. 注意一个点的海温重建可代表多大的范围其实是不确定的. 如果单个重建数据可以影响全球, 这将导致过度校正. 鲁丹期共采用了5个记录点(蓝点), 其中左下方两个点几乎重合. 等值线显示模拟的440Ma表面温度(°C)(Li等, 2022)

滤波, 模拟计算中每一时步都进行集合卡尔曼滤波)来解决.

不难看出, 古气候数据同化的不确定性来源于多个因素, 包括地质记录数量和时空分辨率的影响、古气候代用指标模型的不确定性以及数值模拟本身的偏差. 值得注意的是, 特定地质时期重建记录极度匮乏与空间覆盖的强异质性, 将导致数据-模型协同优化面临本质性困难; 这是制约古气候数据同化可靠性的关键瓶颈. 当前研究尚无法确定支撑稳健古气候数据同化结果所需的最小重建数量和模型模拟成员数量. 一般来说, 新生代以来气候代用记录相对丰富且时空连续性较好, 其同化结果具有较高的可信度; 而更古老的地质时期的同化结果要谨慎看待.

虽然古气候数据同化的不确定性难以避免, 但它体现了地球科学研究方法的转变, 为探索地球过去气候变化提供了一个有力工具. 地球科学正从单一学科的研究转向多学科交叉研究, 从定性研究转向量化过程研究, 从相对独立的数据观测和数值模拟研究转向以新数据获取和大数据支撑的“数据-模式驱动科学研究体系”(郭正堂, 2019), 古气候数据同化恰恰体现了这一转变. 其实, 古气候数据同化的最大优势在于其综合性, 它可以将树轮、冰芯、湖泊沉积物、珊瑚和海洋沉积物等多源数据同时考虑进来, 使得单一数据源的局限性得以克服, 在充足地质证据的约束下, 利用气候模拟形成时间和空间上更加全面、质量更可靠的古气候重建结果. 未来, 深度学习可以建立更好的气候代用指标与气候变量之间的转换模型, 在线同化

系统可以极大提升古气候数据同化的动力学约束, 古气候数据同化结果的可靠性有望进一步提升.

致谢 感谢中国地质大学(武汉)环境学院大气科学系的张泽和王锡津帮忙绘制了图件草图.

参考文献

- 郭正堂. 2019. 《地球系统与演变》: 未来地球科学的脉络. 科学通报, 64: 883–884
- 方苗, 李新. 2016. 古气候数据同化: 缘起、进展与展望. 中国科学: 地球科学, 46: 1076–1086
- 张昊勋, 李明松, 胡永云. 2025. 古气候数据同化: 原理和展望. 中国科学: 地球科学, 55: 416–434
- Erb M P, McKay N P, Steiger N, Dee S, Hancock C, Ivanovic R F, Gregoire L J, Valdes P. 2022. Reconstructing Holocene temperatures in time and space using paleoclimate data assimilation. *Clim Past*, 18: 2599–2629
- Fang M, Li X, Chen H W, Chen D. 2022. Arctic amplification modulated by Atlantic Multidecadal Oscillation and greenhouse forcing on multidecadal to century scales. *Nat Commun*, 13: 1865
- Goosse H, Crespin E, de Montety A, Mann M E, Renssen H, Timmermann A. 2010. Reconstructing surface temperature changes over the past 600 years using climate model simulations with data assimilation. *J Geophys Res*, 115: 2009JD012737
- Judd E J, Tierney J E, Huber B T, Wing S L, Lunt D J, Ford H L, Inglis G N, McClymont E L, O'Brien C L, Rattanasriampaipong R, Si W, Staitis M L, Thirumalai K, Anagnostou E, Cramwinckel M J, Dawson R R, Evans D, Gray W R, Grossman E L, Henahan M J, Hupp B N, MacLeod K G, O'Connor L K, Sánchez Montes M L, Song H, Zhang Y G. 2022. The PhanSST global database of Phanerozoic sea surface temperature proxy data. *Sci Data*, 9: 753
- Judd E J, Tierney J E, Lunt D J, Montañez I P, Huber B T, Wing S L, Valdes P J. 2024. A 485-million-year history of Earth's surface temperature. *Science*, 385: eadk3705
- Li M, Kump L R, Ridgwell A, Tierney J E, Hakim G J, Malevich S B, Poulsen C J, Tardif R, Zhang H, Zhu J. 2024. Coupled decline in ocean pH and carbonate saturation during the Palaeocene-Eocene Thermal Maximum. *Nat Geosci*, 17: 1299–1305
- Li X, Hu Y, Guo J, Lan J, Lin Q, Bao X, Yuan S, Wei M, Li Z, Man K, Yin Z, Han J, Zhang J, Zhu C, Zhao Z, Liu Y, Yang J, Nie J. 2022. A high-resolution climate simulation dataset for the past 540 million years. *Sci Data*, 9: 371
- Lyu Z, Vuille M, Goosse H, Orrison R, Novello V F, Cruz F W, Strikis N M, Cauhy J. 2024. South American monsoon intensification during the last millennium driven by joint Pacific and Atlantic

- forcing. *Sci Adv*, 10: eado9543
- Shi F, Goosse H, Klein F, Zhao S, Liu T, Guo Z. 2019. Monopole mode of precipitation in East Asia modulated by the South China Sea Over the last four centuries. *Geophys Res Lett*, 46: 14713–14722
- Tardif R, Hakim G J, Perkins W A, Horlick K A, Erb M P, Emile-Geay J, Anderson D M, Steig E J, Noone D. 2019. Last Millennium Reanalysis with an expanded proxy database and seasonal proxy modeling. *Clim Past*, 15: 1251–1273
- Tierney J E, Zhu J, King J, Malevich S B, Hakim G J, Poulsen C J. 2020. Glacial cooling and climate sensitivity revisited. *Nature*, 584: 569–573
- Tierney J E, Zhu J, Li M, Ridgwell A, Hakim G J, Poulsen C J, Whiteford R D M, Rae J W B, Kump L R. 2022. Spatial patterns of climate change across the Paleocene–Eocene Thermal Maximum. *Proc Natl Acad Sci USA*, 119: e2205326119
- Valler V, Franke J, Brugnara Y, Brönnimann S. 2022. An updated global atmospheric paleo-reanalysis covering the last 400 years. *Geosci Data J*, 9: 89–107
- von Storch H, Cubasch U, Gonzalez-Rouco J, Jones J, Voss R, Widmann M, Zorita E. 2000. Combining paleoclimatic evidence and GCMs by means of data assimilation through upscaling and nudging (DATUN). In: *Proceedings of the 11th Symposium on Global Climate Change Studies*. GKSS Forschungszentrum Geesthacht GmbH-Publications-E, American Meteorological Society, Long Beach, CA, USA. 28–31
- Zhang Q, Liu B, Li S, Zhou T. 2023. Understanding models' global sea surface temperature bias in mean state: From CMIP5 to CMIP6. *Geophys Res Lett*, 50: e2022GL100888

(责任编辑: 岳天祥)