

差分进化算法在气候变化中的应用

王宁^{1*}, 刘俊宏^{2*}

¹陕西省气象信息中心, 卫星遥感备份业务科, 陕西 西安

²黑龙江省黑河市气象局, 综合保障中心, 黑龙江 黑河

收稿日期: 2024年4月17日; 录用日期: 2024年5月15日; 发布日期: 2024年5月22日

摘要

目前, 气象预测依赖于先进的气象观测技术、大数据分析和计算机模型等手段, 以提供更为准确的天气预报和气候预测。尽管通过不断改进算法和提高数据采集质量, 气象预测的准确性和时效性得到了显著提升, 但在数据处理、模型训练和特征选择等方面仍存在一些不足之处。引入差分进化算法(Differential Evolution, DE)不但可以弥补这些不足而且也能提升气候科学研究的创新水平。本文重点关注DE算法在优化气候模型性能、特征选择和数据处理方面的作用。通过该算法进行气候模型参数优化, 改进模型的拟合能力和预测准确性, 从而提高气候变化预测的可靠性。DE算法的引入将为气象数据处理和气候科学研究提供新的有效途径, 有望促进气候变化预测的准确性和可靠性, 推动气候科学领域的创新发展。

关键词

DE算法, 模型优化, 气候变化

Application of Differential Evolution Algorithm in Climate Change

Ning Wang^{1*}, Junhong Liu^{2*}

¹Satellite Remote Sensing Backup Business Department, Shaanxi Provincial Meteorological Information Center
Xi'an Shaanxi

²Comprehensive Support Center, Meteorological Bureau of Heihe City, Heilongjiang Province, Heihe,
Heilongjiang

Received: Apr. 17th, 2024; accepted: May 15th, 2024; published: May 22nd, 2024

Abstract

At present, meteorological forecasting relies on advanced meteorological observation techniques,

*共同第一作者。

文章引用: 王宁, 刘俊宏. 差分进化算法在气候变化中的应用[J]. 计算机科学与应用, 2024, 14(5): 68-74.

DOI: 10.12677/csa.2024.145115

big data analysis, and computer models to provide more accurate weather and climate forecasts. Although the accuracy and timeliness of meteorological prediction have been significantly improved by continuously improving algorithms and improving data collection quality, there are still some shortcomings in data processing, model training, and feature selection. The introduction of the Differential Evolution (DE) algorithm can not only compensate for these shortcomings but also enhance the innovation level of climate science research. This article focuses on the role of the DE algorithm in optimizing climate model performance, feature selection, and data processing. By using this algorithm to optimize climate model parameters, the fitting ability and prediction accuracy of the model are improved, thereby enhancing the reliability of climate change prediction. The introduction of the DE algorithm will provide a new and effective way for meteorological data processing and climate science research, which is expected to promote the accuracy and reliability of climate change prediction, and promote innovative development in the field of climate science.

Keywords

DE Algorithm, Model Optimization, Climate Change

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

天气条件是影响农业生产的重要因素之一。随着全球气候变暖和人类活动的影响,天气条件的变化趋势和规律也变得日益复杂和多样化[1]。我国是一个农业大国每年都会受到天气条件的严重影响。2023年5月17日,石家庄多地突降冰雹,致使农户的蔬菜大棚遭到严重的破坏,6月12日北京密云遭遇特大冰雹灾害,短短十几分钟村里137座蔬菜大棚约200亩蔬菜作物严重受损。同年7月17日至24日,河南出现了罕见的极端降雨天气,新乡、周口、开封、安阳、焦作、鹤壁、郑州等多个地区均受到了不同程度的洪涝灾害,部分地区由于河道水位较高,导致农田积水,粮食减产,生猪等畜禽及圈舍直接损失高达22.5亿元。

气候变化是一个极为复杂的过程,受到自然变化和人类活动等多种因素的影响。在这种情况下,气象数据的获取和质量直接影响着气候变化预测的准确性。然而,由于观测站点有限、数据收集方法不统一等原因,气候数据常常存在着不完整和不准确的情况,从而影响了预测结果的可靠性。此外,气候系统的复杂性也使得建立准确的气候模型变得异常困难,当前的气候模型虽然能捕捉部分气候变化特征,但却难以完全模拟整个系统的复杂行为,因而导致预测结果存在一定的局限性。

尽管目前气象数据处理已经取得了一定的成绩,但仍然存在一些挑战和不足之处,需要进一步改进和完善。未来的研究方向包括改进数据质控制和处理方法、提高数据处理效率和算法性能、优化模型选择和评估方法、改进特征选择和工程方法,以及提高实时性和预测性等方面。

气象数据常受到噪声和异常值的干扰,处理不当直接影响模型训练和预测的结果。因此,改进数据质量控制和处理方法是一项重要挑战。针对不同类型的气象现象(如温度、湿度、降水等),选择合适的模型进行预测是必要的。气象数据具有大维度的特征,但并非所有特征对预测目标都有用,有效的特征选择和工程方法能够提高模型性能,但如何选择合适的特征和进行有效的特征工程仍没有统一的标准。

本文旨在针对气候变化受多重因素影响的问题,采用DE算法来优化气候模型参数,改善特征选择,

增强数据处理能力, 并优化模型结构和参数。这一举措旨在提升气候模型的预测能力, 为简化气候模型结构、提高预测准确性提供有力支持, 为深入理解气候变化规律和趋势提供新的思路 and 工具。通过这些努力, 有望促进气候科学研究的发展与进步, 进而协助科学家和政策制定者更好地理解 and 应对气候变化所带来的挑战。

2. 现状

2.1. 气温和湿度

目前, 气温和湿度参数优化面临着数据不准确和复杂性的挑战。由于气候系统受到诸多因素影响, 如大气运动、海洋循环和地形, 准确建立双方的关联并进行优化变得十分复杂。此外, 数据质量和获取也是问题所在, 传感器误差、数据缺失和不一致性影响模型优化。为了解决这些挑战, 可采取以下方法: 利用更多数据源、模型集成优化、特征工程和模型调参等方法, 通过持续改进建模方法、数据质量和模型优化策略, 逐步提高气候变化检测的准确性和可靠性。

2.2. 风力的特征选择

为了解决气象中风力特征选择的不足, 需要针对特征相关性高、关键特征缺失、特征工程不足和数据质量问题采取有效措施。具体可通过进行全面的特征分析、利用领域知识和先进技术进行特征挖掘、确保数据质量的严格管控以及持续改进特征工程方法来提高对风力变化的预测准确性和可靠性, 以适应气象领域的复杂性和多变性。

2.3. 降雨的数据处理

气象中降雨数据处理存在数据缺失和不准确、时间和空间分辨率不足以及数据质量问题。为了解决这些问题, 可采用插补方法填补数据缺失, 利用先进的遥感技术和气象雷达等设备获取更高分辨率的降雨数据, 并引入先进的数据处理和质量评估技术。这些改进方法将提高对降雨的预测能力, 更好地应对气象灾害和水资源管理等相关问题。

2.4. 雨和冰雹模型优化

气象中关于雨和冰雹模型的优化存在参数化方案不完善、物理过程模拟精度有限以及数据源和观测缺乏等问题。为了解决这些问题, 可通过改进参数化方案、加强对物理过程的模拟研究以及加强数据收集和观测网络建设来提升模型的准确性和可靠性。这些优化措施将为气象预报和灾害防范提供更有效的支持。

3. 气候变化的基本规律

3.1. 温度和湿度的变化趋势和规律

随着全球气候变暖, 温度和湿度的变化趋势和规律也日益复杂和多样化, 许多地区夏季温度越来越高, 冬季温度则越来越低[2]。这些显著变化对农业生产和发展带来了巨大挑战。地表温度、土壤湿度和大气湿度三者相互关联, 在作物生长周期中扮演着重要角色。以 0°C 为分界点, 当温度高于 0°C 时, 温度与土壤湿度呈正相关, 地表温度会下降; 反之, 当温度低于 0°C 时, 土壤湿度的变化则会对地表温度产生相反的影响。种子只有在适宜的条件下才能发芽生长, 而长时间的低温会影响种子的存活率。

3.2. 风的变化趋势和规律

风的产生是由太阳的照射致使地表空气受热膨胀上升, 形成了热空气。与此同时, 冷空气则在地表

横流。上升中的热空气逐渐冷却下沉与地表的热空气相互接触, 如此反复形成了风。风力的大小主要受温差、气压带与风带、气候变化以及地形等多种因素的影响[3]。在温差较大的地区, 水平气压梯度力增大导致风力增强。受风带控制时, 通常会出现较强的风力, 尤其是在西风带。相反, 在受高气压带控制下, 风力往往较弱, 甚至可能没有风, 特别是在副热带高气压带。此外, 地形起伏较缓且受迎风坡作用的地区一般风力较大, 而背风坡和地表崎岖的地区风力相对较小。

3.3. 降雨的变化趋势和规律

降水量受到诸多因素的影响, 主要受地形地势和气候带的影响较大。目前没有统一的标准来评估所有地区的降雨形势, 需要根据具体情况做出相应的调整。降雨的核心条件包括水汽供应、达到饱和状态的上升气流, 以及促使水汽凝结成云和降水的凝结核。在分析降水时, 需要综合考虑降雨的核心条件和本地的主导因素, 这有助于提高对降雨规律的准确性。

针对不同地区的降雨情况, 可以根据当地的地形地势和气候特点来进行调整和分析。例如, 对于山区地形, 可能会受到地形抬升引起的降水影响; 而对于沿海地区, 则可能更多受到海气的湿润影响。因此, 在研究降水规律时, 需要考虑到地方特有的因素, 以便更准确地理解和预测降水情况。综合考虑降雨的核心条件和当地主导因素可以帮助提升对降雨规律的把握, 为气象预测、水资源管理等提供更可靠的依据。

4. DE 算法

4.1. DE 算法介绍

20世纪90年代由美国学者 Rainer Storn 和 Kenneth Price 在遗传算法(Genetic Algorithm, GA)等进化思想的基础上提出差分进化算法(Differential Evolution, DE) [4], 本质是一种多目标(连续变量)优化算法(MOEA), 主要用于多维空间寻找整体最优解。DE 算法又称作智能优化算法和启发式算法, 如人工神经网络(Artificial Neural Networks, ANN), 文化算法(Cultural Algorithm, CA)等都属于启发式的优化算法。通过模拟遗传学中的杂交(crossover)、变异(mutation)、复制(reproduction)来设计遗传算子。

差分进化算法相对于遗传算法而言, 相同点都是通过随机生成初始种群来开始优化过程。在优化过程中, 种群中每个个体的适应度值被用作选择的标准, 以便确定哪些个体将被用于产生下一代[5]。主要过程包括变异、交叉和选择, 变异操作用于引入种群的多样性, 交叉操作用于产生新个体, 选择操作用于确定哪些个体将在下一代中生存或繁殖[6]。尽管 DE 和 GA 有一些相同的操作步骤, 但它们在具体的变异和选择策略上存在一些显著的差异, 这也导致了它们在优化效果和收敛速度等方面的差异。DE 的变异操作通过计算种群中不同个体间的差值并缩放后加到其他个体上, 以此来生成新的候选解, 这种方法能够有效地避免局部最优并保持种群的多样性[7]。而 GA 的变异操作通常是对个体的基因进行随机改变, 可能会产生与现有个体相似或完全不同的新个体。

4.2. DE 算法提升气候预测能力

4.2.1. 参数优化

差分进化算法作为一种元启发式优化算法, 在近年来在优化问题中得到了广泛的应用。在气候科学领域, DE 算法被用于调整气候模型中的参数, 以提高模型的准确性和可靠性。这种参数优化过程是复杂的, 需要综合考虑算法的特性、问题的性质以及可用的计算资源。在对气象资料进行参数优化时, 需要尽可能综合参数自适应调整策略和参数的静态设定。通过优化模型参数, 可以更好地拟合观测数据, 从而有效提升气候变化预测的精度。这种方法有助于克服现有气候模型中存在的局限性, 特别是在捕捉复杂的气候系统行为方面。因此, DE 算法在气候科学中的应用为改进气候模型提供了一种有效的手段, 它

能够通过优化模型参数来提高模型的适应性, 进而增强气候变化预测的准确性和可靠性。

4.2.2. 特征选择

差分进化算法主要用于解决优化问题, 包括连续变量和离散变量的优化。在机器学习中, 特征选择是提高模型性能的重要步骤, 它可以帮助减少模型的复杂度、提高训练速度, 并降低过拟合的风险。虽然差分进化算法本身并非专门为特征选择而设计, 但它仍然可以应用于特征选择任务, 尤其是针对连续特征选择问题。

特征选择问题可以转化为适合差分进化算法求解的优化问题。通过合理定义适应度函数或目标函数, 并利用差分进化算法的全局搜索能力, 可以有效地找到最优的特征子集, 从而提高模型的性能。值得注意的是, 特征选择的效果取决于所使用的机器学习模型和数据集的特性, 因此在实际应用中可能需要进行多次实验和调整。

利用差分进化算法能够帮助选择最具代表性的气候特征, 从而简化气候模型的复杂性并提高预测效果。特征选择在气候科学研究中起着至关重要的作用, 它能够剔除对气候变化预测影响较小或不重要的特征, 从而降低模型的复杂度, 提升模型的泛化能力。通过精心选择气候特征, 能更好地理解气候系统中的关键因素, 并减少模型在处理过多信息时出现的过拟合问题。

4.2.3. 数据处理

气候数据经常会受到各种来源的干扰, 例如噪声、缺失值等问题会影响数据的准确性和完整性, 进而对气候预测模型的建立和应用造成不利影响。DE 算法通过优化数据处理方法, 帮助消除或减弱气候数据中的噪声干扰, 填补缺失值, 甚至校正数据中的错误, 从而提高数据的质量和可靠性。通过对数据进行优化处理, 更好地反映真实的气候变化趋势和规律, 为气候预测模型提供更加可靠的输入。

4.2.4. 模型优化

差分进化算法在结合机器学习模型和深度学习模型进行气候变化预测方面展现出了巨大的潜力。气候数据通常具有复杂的特征和变化规律, 传统的建模方法无法充分捕捉其中的关联性和趋势, 导致气候变化预测的准确性和稳定性不尽如人意。

通过 DE 算法对机器学习模型和深度学习模型进行参数优化, 帮助模型更好地拟合气候数据, 提高模型的泛化能力和预测准确性。DE 算法能够有效地搜索参数空间, 快速找到最优解, 使得模型在处理气候数据时能够更加有效地学习和预测。同时, 结合 DE 算法的优化能力, 可以避免模型过拟合或欠拟合的问题, 提升模型的整体性能和稳定性。

4.3. 算法步骤

4.3.1. 初始化种群

差分进化算法初始化种群:

$$\{X_i(0) | x_{i,j}^L \leq x_{i,j}(0) \leq x_{i,j}^U; i=1,2,\dots,NP; j=1,2,\dots,D\} \quad (1)$$

其中, $x_i(0)$ 是第 i 个个体, j 表示第 j 维。

$$x_{i,j}(0) = x_{i,j}^L + \text{rand}(0,1)(x_{i,j}^U - x_{i,j}^L) \quad (2)$$

其中, $x_{i,j}^L$ 和 $x_{i,j}^U$ 分别为第 j 维的下界和上界, $\text{rand}(0,1)$ 表示在区间 $[0, 1]$ 上的随机数。

4.3.2. 变异

DE 算法利用差分策略进行个体变异, 一种常见的差分策略是从种群中随机选择两个不同的个体, 计

算它们的向量差, 对该差异向量进行缩放后与待变异个体进行向量合成。

$$V_i(g+1) = X_{r_1}(g) + F(X_{r_2}(g) - X_{r_3}(g)) \quad (3)$$

其中, r_1 , r_2 和 r_3 是三个随机数, 区间为 $[1, NP]$, F 称为缩放因子, 为一个确定的常数, NP 为种群规模。 g 表示第 g 代。

4.3.3. 交叉

交叉操作的目的是随机选择个体, 差分进化算法本身也是一种基于随机性的算法, 交叉操作的方法是:

$$U_{i,j}(g+1) = \begin{cases} V_{i,j}(g+1) & \text{if } \text{rand}(0,1) \leq CR \\ x_{i,j}(g) & \text{otherwise} \end{cases} \quad (4)$$

其中, CR 称为交叉概率。通过概率的方式随机生成新的个体。

4.3.4. 选择

在差分进化算法中采用的是贪婪选择策略, 即选择较优秀的个体作为新一代的个体。

$$X_i(g+1) = \begin{cases} U_i(g+1) & \text{iff } (U_i(g+1)) \leq f(X_i(g)) \\ X_i(g) & \end{cases} \quad (5)$$

DE 算法的群体演化过程由突变和选择驱动。突变阶段包括突变和交叉操作, 这些操作旨在利用或探索搜索空间, 而选择阶段则确保有潜力的个体信息能够被进一步利用。DE 算法流程图见图 1。

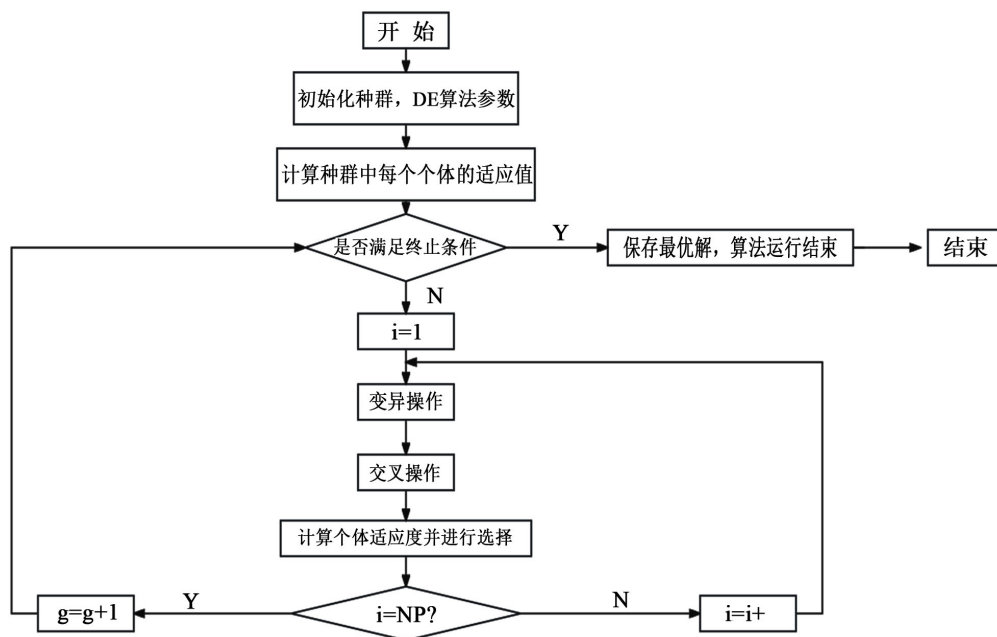


Figure 1. DE algorithm flowchart

图 1. DE 算法流程图

4.4. 结论

差分进化算法(DE 算法)是一种基于群体搜索和优化的算法, 常用于解决优化问题。在气象数据处理、

模型训练和特征选择方面, DE 算法相对传统方法可能具有以下优势:

1. 全局搜索能力强: DE 算法具有强大的全局搜索能力。通过迭代更新种群中的个体, 能够更好地发现目标函数的全局最优解。在数据处理和模型训练中, 利用 DE 算法能够有效地搜索参数空间, 找到更优的参数组合, 提高模型的性能和预测准确性。其全局搜索能力使得 DE 算法在处理复杂的气象数据和模型优化中具有重要意义, 有助于更好地理解气候变化规律和趋势, 为科学家和政策制定者提供更可靠的数据支持和决策依据。

2. 免受初始点影响: DE 算法不受初始点选择的影响, 这是其一个重要优点。相比于一些传统的优化算法, DE 算法不需要事先给定初始解, 而是通过随机生成初始种群, 并在搜索过程中自适应地调整参数。因此, DE 算法对初始点的选择不敏感, 这种特性使得在处理数据和训练模型时更加稳健。DE 算法能够更可靠地找到全局最优解, 避免局部最优解的陷阱, 在实际应用中更加灵活和可靠, 适用于各种气象数据处理和模型优化的场景。

3. 并行计算能力: DE 算法具有良好的并行计算能力, 这使得它能够在多处理器系统或分布式计算环境下进行高效计算。通过并行计算, DE 算法可以同时多个个体进行搜索和评估, 加速了搜索过程特别是在处理大规模数据和复杂模型训练时具有明显优势。在数据处理方面, DE 算法还可并行地对数据进行分割和处理, 每个处理单元独立地进行计算最后合并结果, 提高了处理数据的速度和效率。这对于处理大规模气象数据、气候模拟和预测等任务非常重要。

5. 总结

气候变化对农作物的生长周期产生重要影响, 其中最活跃的因素包括温度、湿度、降水、霜冻和冰雹。了解气象要素的生成条件和变化规律, 能够帮助及时采取有效措施, 为农业生产发展和高效管理提供合理的科学依据和实践指导。

本文利用差分进化算法(DE 算法)来解决气候变化受多重因素影响的问题。通过优化气象模型参数、改善特征选择和增强数据处理能力以及优化模型结构和参数提高气象预测准确性, 为气候研究提供新思路。该算法在气候模型参数优化和特征选择方面的应用, 能够有效地搜索参数空间和特征子集, 找到最优解或最佳特征组合, 推动气候预测领域取得更为显著的进展。

参考文献

- [1] 张卓珺. 干旱气候对农业生产的影响及气象服务对策[J]. 农业灾害研究, 2022, 12(2): 128-130.
- [2] 贺晨昕. 基于气候变化的农业气象灾害对春玉米产量的影响[J]. 种子科技, 2023, 41(8): 133-135.
- [3] 王利霞. 气候变化对农业气象灾害与病虫害的影响分析[J]. 新农业, 2022(7): 13.
- [4] 宋贤芳, 杨扬, 张勇, 等. 强化学习引导的产品变更路径多目标差分进化算法[J/OL]. 控制理论与应用, 1-9. <http://kns.cnki.net/kcms/detail/44.1240.TP.20240229.1733.022.html>, 2024-04-08.
- [5] 徐王颖, 于小兵. 基于多自适应算子的改进差分进化算法[J]. 信息技术, 2024(2): 22-30+38. <https://doi.org/10.13274/j.cnki.hdzj.2024.02.004>
- [6] 王冠中, 王士军, 冉川东. 基于改进差分进化算法的自由曲面测量路径优化[J]. 制造技术与机床, 2024(3): 51-56. <https://doi.org/10.19287/j.mtmt.1005-2402.2024.03.007>
- [7] 黄亚伟, 钱雪忠, 宋威. 基于双档案种群大小自适应方法的改进差分进化算法[J/OL]. 计算机应用, 1-14. <http://kns.cnki.net/kcms/detail/51.1307.TP.20240305.0850.002.html>, 2024-03-29.