

一种优于动力次季节温度预测的机器学习模型

薛翔海^{1,2}

¹湖南师范大学地理科学学院, 湖南 长沙

²地理空间大数据挖掘与应用湖南省重点实验室, 湖南 长沙

收稿日期: 2025年3月16日; 录用日期: 2025年4月8日; 发布日期: 2025年4月15日

摘要

可靠的次季节温度预测对极端温度事件的防灾减灾至关重要。然而, 现有的次季节温度预测动力模型常受到初值问题和边值问题的影响, 导致其预报能力相对薄弱。尽管近年来机器学习模型在次季节预测中逐渐展示出超越动力模型的潜力, 但中国次季节温度预测仍主要依赖于动力学模型。鉴于此, 本研究基于Lasso (Multi-task Lasso)机器学习算法, 构建了覆盖中国所有格点的次季节温度预测模型, 并采用余弦相似度指标评估Lasso和CFSv2 (The Climate Forecast System version 2)动力模型在2018~2022年测试期内的预测性能表现。结果表明: Lasso在整体预测技能上显著优于CFSv2, 其在未来3~4周和5~6周的平均余弦相似度较CFSv2分别提升了0.33和0.34; 并且, 在常规温度情景下, Lasso能够更精准地捕捉温度变化的规律, 80%以上月份的平均CS高于CFSv2; 其仅在极端低温情景下存在一定局限性, 预测技能略逊于CFSv2。

关键词

动力模型, 机器学习模型, 次季节温度预测, 中国

A Machine Learning Model Superior to Dynamic Subseasonal Temperature Forecasting

Xianghai Xue^{1,2}

¹School of Geographical Sciences, Hunan Normal University, Changsha Hunan

²Key Laboratory of Geospatial Big Data Mining and Application, Changsha Hunan

Received: Mar. 16th, 2025; accepted: Apr. 8th, 2025; published: Apr. 15th, 2025

文章引用: 薛翔海. 一种优于动力次季节温度预测的机器学习模型[J]. 数据挖掘, 2025, 15(2): 176-183.
DOI: 10.12677/hjdm.2025.152015

Abstract

Reliable subseasonal temperature forecasting plays an important part in extreme temperature events prevention and mitigation. However, current dynamical models for subseasonal temperature forecasting are often influenced by initial value and boundary value problems, resulting in relatively weak forecasting performance. Although machine learning models have shown potential in surpassing dynamical models for subseasonal forecasting in recent years, subseasonal temperature forecasting in China still mainly relies on dynamical models. Under this background, the study constructs a subseasonal temperature forecasting model covering 957 grid points across China based on the Lasso (Multi-task Lasso) machine learning algorithm and uses the cosine similarity metric to evaluate the performance between the Lasso and CFSv2 (The Climate Forecast System version 2) dynamic model during the test period from 2018 to 2022. The results show that the Lasso significantly outperforms CFSv2 in overall forecasting performance. The average cosine similarity of the Lasso is 0.33 and 0.34 higher than the CFSv2 at the forecast horizon of weeks 3~4 and 5~6, respectively. Moreover, in normal temperature scenarios, the Lasso can more accurately capture temperature variation patterns with the average cosine similarity for over 80% of the months higher than that of the CFSv2. However, the Lasso has some limitations in forecasting extreme low temperature scenarios, where its forecasting skill is slightly inferior to that of the CFSv2.

Keywords

Dynamic Model, Machine Learning Model, Subseasonal Temperature Forecasting, China

Copyright © 2025 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

全球变暖的背景下，寒潮和热浪等极端温度事件发生的频率显著增加[1]。此类事件不仅会造成巨大的社会经济损失[2]，还会对人类健康造成影响[3]。提前期介于2周到两个月之间的次季节温度预测能为各行业和社会应对极端温度事件提供充分的准备时间，从而减轻此类事件带来的危害和损失[4]。中国作为世界上人口最多的国家之一，频发的极端温度事件更容易对社会经济发展和自然生态环境造成严重的影响[5]。因此，构建可靠的中国次季节温度预测对于减轻极端冷热事件所造成的危害至关重要。

尽管在过去几十年中，一周内的天气预报，以及两个月至九个月的气候预测都取得了显著进展[6]-[10]，但提前期为两周至两个月的次季节预测研究仍处于初级阶段。为了填补短期天气预报到长期气候预测之间的缝隙，2014年由世界天气研究计划和世界气候研究计划共同发起了次季节-季节研究计划，以期提高两周到两个月时间尺度的预测精度[4]。目前，欧洲中期天气预报中心(European Centre for Medium-Range Weather Forecasts, ECMWF)、美国国家环境预报中心(National Centers for Environmental Prediction, NCEP)、英国气象局(Met Office)等都发布各自的季节-次季节模型输出数据。然而，次季节预测位于具有初值问题的短期天气预报与具有边值问题的长期气候预测二者之间[11]，常同时面临两类问题的挑战。一方面，初始误差随时间迅速增长，在超过两周后，达到最大可预测时效，误差的增大致使预报失去参考价值；另一方面，大气虽然开始受到外界强迫影响，但尚未完全被外界条件所主导[12]。初边值问题的交织导致当今的业务动力数值模式对次季节的预报能力相对薄弱。此外，由于大气的混沌性、非线性和复杂性，导致气温变化的物理过程尚不完全清晰，通过动力模型模拟气温变化实际物理过程存在显著局

限性。因此,基于上述原因,动力模型在次季节温度业务预报中的实际效用仍有待进一步验证[13]。

近年来随着人工智能的发展,数据驱动模型在气象预测上取得了显著的进展[9] [14] [15],此类模型通过数据统计或者机器学习等方法,利用历史数据去预测未来气象变量,展现出较高的预测精度和效率。此外,多项研究表明,厄尔尼诺与南方涛动(ENSO)、马登-朱利安振荡(MJO)、土壤湿度的大尺度异常和位势高度等是次季节时间尺度上重要的可预测性来源[16] [17],但这些可预测性来源在现有天气和气候模型中的表征仍不够完善,因此,利用数据驱动方法来建模这些重要预测来源与次季节温度之间的关系,为进一步改进次季节预测提供了可能[16]。目前,数据驱动模型已经被应用于美国[13]、泰国[18]和欧洲中部[19]的次季节变量预测,并被证明具有可行性。He 等人的研究进一步指出[20],数据驱动的机器学习模型 Lasso 在美国次季节温度预测中的整体预测性能显著优于动力模型。然而,中国次季节温度预测仍主要依赖于动力模型。针对这一现状,本研究将 Lasso 算法引入中国次季节温度预测任务,构建了覆盖中国所有格点的次季节温度预测模型,并选择了 CFSv2 动力模型与 Lasso 进行对比,以探究 Lasso 在中国次季节温度预测中性能优势和局限性,从而为提升中国次季节温度预测能力提供科学依据和技术支持。

2. 研究区概况与数据

2.1. 研究区概况

中国地域辽阔,气候多样,温度特征因地域差异而显著不同,涵盖了热带,亚热带,温带和寒带等多种气候类型。而且,中国地处东亚季风区,在季风影响下极端温度事件频发[21],其中,北方地区是寒潮易发区,南方地区是夏季热浪高发区[22]。

2.2. 数据描述

本研究参考 He 等人为美国次季节预测构建数据集的方法[13],统一格式整合了多源数据的多个变量,构建了适用于中国次季节温度预测的数据集。该数据集可以被分为时空数据和时间数据两类,以下分别介绍了这两类数据的组成和来源。

对于时空数据,两米处的日气温(tm)、日降水(precip)和月土壤湿度(sm)数据分别来自于 NOAA 的 Global Unified Temperature 数据集、Soil Moisture V2 数据集和 Unified Gauge-Based Analysis of Daily Precipitation 数据集(<https://psl.noaa.gov/>)。这些数据的空间覆盖范围与研究区保持一致。此外,日海平面气压(slp)、日相对湿度(rhum)、日 10 hpa 和 500 hpa 高度的位势高度(hgt10 和 hgt500)数据来源于 NOAA 的 National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) Reanalysis 1 数据集(<https://psl.noaa.gov/>)。这些数据的空间覆盖范围为赤道至北纬 60°、东经 70°至 250°。小时的海表温度(sst)数据来源于 Medium-Range Weather Forecast Reanalysis 5 (ERA5),由气候数据存储平台提供(<https://cds.climate.copernicus.eu/>),覆盖范围为南纬 20°至北纬 65°、东经 150°至 250°。

对于时间数据,于日 MJO 相位和振幅[23],以及双月的多变量厄尔尼诺南方涛动指数(MEI) [24]分别来自于澳大利亚气象局(<http://www.bom.gov.au/>)和美国国家海洋和大气管理局地球系统研究实验室(<http://pls.noaa.gov/enso/mei/>)。此外,每周的 Niño 1 + 2, 3, 3.4 [25]和日北大西洋震荡指数(NAO) [26]来自美国国家气象局(<https://www.cpc.ncep.noaa.gov/data/>)。

2.3. 数据预处理

首先,为统一数据的时间分辨率,本研究将每周、每月和每两个月的数据拆分为每日数据,并将每小时的数据平均为每日数据,从而生成日尺度的数据集。其次,对于预测量日均温度被插值到 1°的经纬度格网中,并转化为随后两周的平均温度。第三,由于时空数据维度较高,大多机器学习模型无法直接使用这

些数据。因此,本研究采用主成分分析(Principal Component Analysis, PCA),根据 1986 年至 2017 年时空变量的主成分载荷,分别提取对应时空变量的主成分(PCs)。在确定各变量的主成分数量时,本研究遵循以下准则:首先计算前五个主成分的累计方差贡献率,若其超过 85%,则选取前五个主成分;否则,逐步增加主成分数量直至累计方差贡献率达到 85%以上。通过该方法,我们初步确定了各变量的主成分数量。然而,在模型训练过程中发现,sm 和 precip 采用前五个主成分会导致模型性能下降。因此,我们基于训练集的预测精度,通过系统性的尝试与调整,最终确定了各变量的最优主成分数量:sm 为 3 个,tm 为 1 个,sst 为 1 个,rhum 为 5 个,slp 为 5 个,hgt10 为 2 个,hgt500 为 5 个,precip 为 4 个。最后,预测量和预测因子的 PCs 都通过每年的每一天进行 Z-score 标准化,而时间变量则直接使用原始数据,不进行标准化处理。

3. 研究方法 with 评价指标

3.1. 机器学习模型

3.1.1. Multitask Lasso

根据 He 等人的研究,机器学习模型在次季节的应用普遍存在过拟合的问题[13]。Multitask Lasso (Lasso)通过一个共同的正则化项进行约束,可以有效减少过拟合,提高模型的泛化能力[27]。He 等人已经证明 Lasso 在美国次季节温度预测任务中的可行性[13],因此本研究使用 Lasso 开展中国次季节温度预测。Lasso 的输入是预测初始日期的时空变量的各个主成分分量以及时间变量,输出是预测初始日期未来第三至四周或第五至六周的两周的平均温度。

3.1.2. 评估准则

本研究为 2018~2022 年测试期内的每个月,分别独立地建立预测模型。为了模拟实时系统,生成了 261 个测试日(2018~2022 年期间,每周 1 天),并根据其年月将其划分为 60 个测试集。每个测试集包含 4 或 5 个测试日。每个测试集的评估流程包括两个部分:(1) 使用滑动窗口策略为时间序列数据创建 5 重训练-验证集,以进行超参数优化。每个验证集的数据与测试集的月份相同,并利用过去五年(2013~2017)的数据形成 5 个验证集。相应的训练集则包括每个验证集之前的 10 年数据。(2) 构建一个用于训练-测试的训练集,其中训练集包含前 24 年的数据。为了避免训练集和测试集之间的重叠,测试集开始的日期必须在训练集结束日期之后 28 天,如图 1 所示。

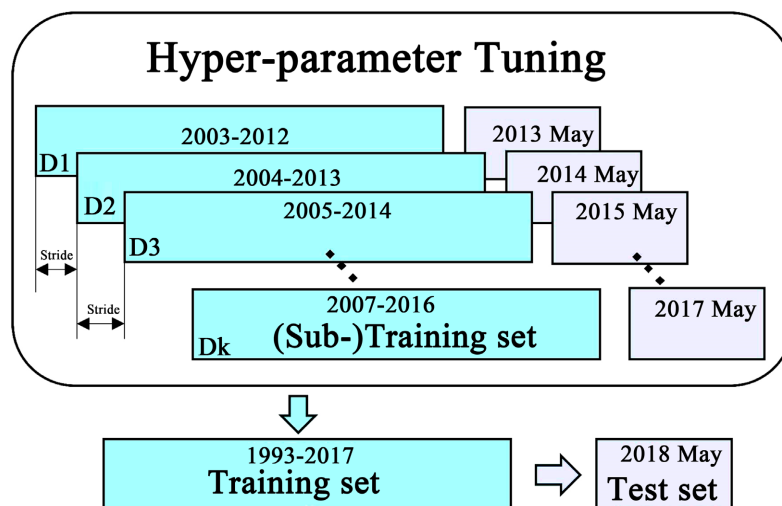


Figure 1. Evaluation pipeline to test in May 2018

图 1. 2018 年 5 月的测试评估准则

3.2. 动力模型

气候预报系统(The Climate Forecast System, CFS)是一个综合模型,集成了海洋、陆地和大气之间的相互作用,由NECP于2004年8月实施。CFS的新版,即CFSv2[28],整合了多个新的物理包,涵盖云-气溶胶辐射、大气地表、海洋和海冰过程,并且在更高的分辨率下运行。CFSv2的温度预测在亚洲夏季季风区域的预测技能与ECMWF最新的模型相当[29]。目前,NCEP将CFSv2作为主要的次季节温度预测产品公开发布。因此,本研究使用CFSv2作为与机器学习模型进行比较的动力模型。

3.3. 评估指标

模型的性能可以通过多种统计指标进行评估。本研究使用余弦相似度(Cosine Similarity, CS)评估模型的预测准确性。CS是次季节气候预测牛仔竞赛中使用的唯一评价指标[30],其定义如下:

$$\cos(\hat{y}, y^*) = \frac{\langle \hat{y}, y^* \rangle}{\|\hat{y}\|^2 \|y^*\|^2} \quad (1)$$

其中, $\hat{y}$ 是测试日预测异常值的向量, y^* 是对应的实际值, $\langle \hat{y}, y^* \rangle$ 表示两个向量的内积, $\|y\|$ 表示向量 y 的模长。CS 的值越接近 1, 代表模型的预测性能越好。

4. 结果与讨论

为了探究 Lasso 和 CFSv2 在中国次季节温度的预测效果,本研究计算了两个模型在测试期间(预测延伸期为第 3~4 周和 5~6 周)所有格点上取得的平均 CS。对于预测延伸期为第 3~4 周, Lasso 取得的平均 CS 为 0.30, CFSv2 为-0.03。对于预测延伸期为 5~6 周, Lasso 取得的平均 CS 为 0.30, CFSv2 为-0.04。基于此,可以得出 Lasso 的整体预测技能显著高于动力模型 CFSv2。为了进一步去细化比较两个模型在预测期的预测能力,图 2 展示了两个模型在全国所有格点预测结果评价指标的占比直方图。对于 Lasso 的预测结果,在预测延伸期为第 3~4 (5~6)周中,有 73.3% (70.7%)的格点 $CS > 0.2$, 31.3% (31.7%)的格点 $CS > 0.4$; 对于 CFSv2 的预测结果,仅有 20.7% (20.3%)的格点 $CS > 0.2$, 7.6% (8.2%)的格点 $CS > 0.4$ 。这一结果表明, Lasso 方法显著提升了中国次季节温度预测精度。

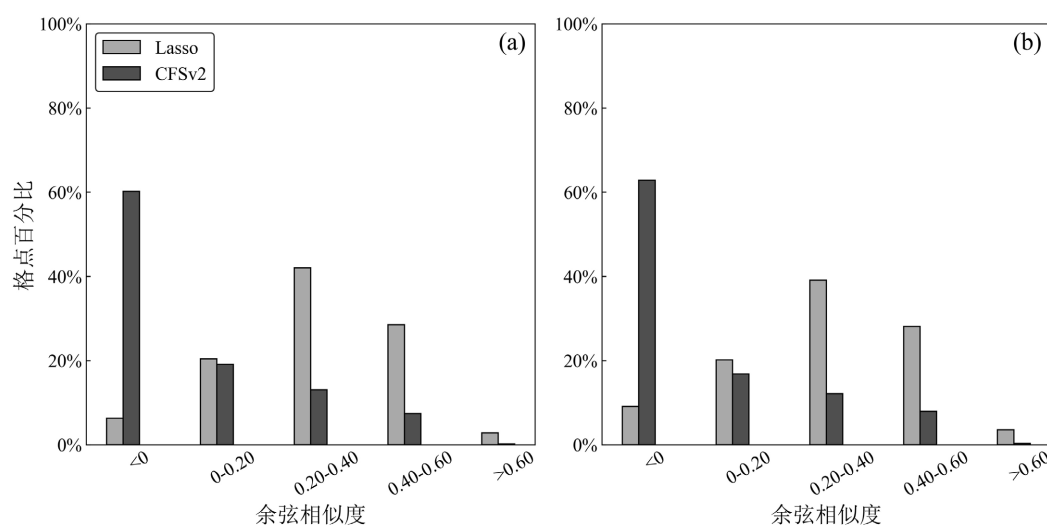


Figure 2. Percentage of grid points for five different ranges of CS values from four two models at the forecast horizons of weeks 3~4 (a) and weeks 5~6 (b)

图 2. 两种模型在预测延伸期为第 3~4 周(a)和第 5~6 周(b)的在五个不同范围的 CS 值所占网格点的百分比

虽然 Lasso 在整体上的预测技能高于 CFSv2, 但由于不同建模模式的模型在不同温度情景下的预测性能有所差异, 因此仅评估模型在整个测试期间的性能是不充分的。基于此, 本研究计算了测试期内每个月, Lasso 和 CFSv2 在所有格点上的平均 CS, 以及两周的温度异常平均值, 并展示在图 3 和图 4, 分别对应预测延伸期为第 3~4 周和第 5~6 周。图 3 显示, 对于预测延伸期为第 3~4 周, Lasso 在 81.7% 月份取得的平均 CS 高于 CFSv2; 图 4 表明, 对于预测延伸期为第 5~6 周, Lasso 在 88.3% 月份取得的平均 CS 高于 CFSv2, 这表明机器学习模型的整体预测能力显著优于动力模型。

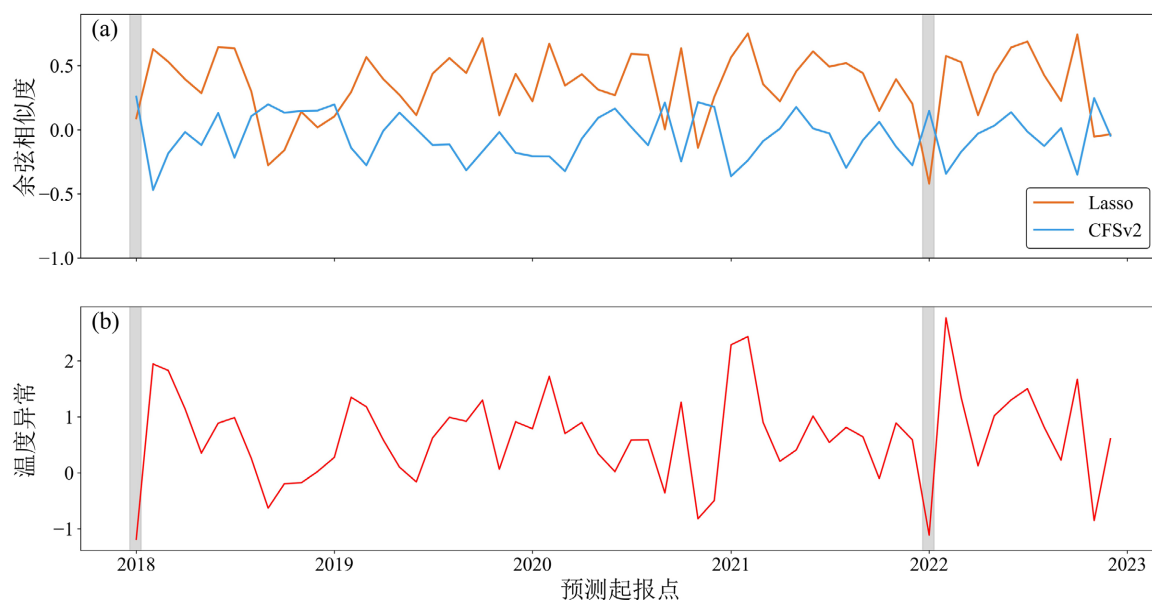


Figure 3. Average CS of two models (a), and average temperature anomalies (b) in each month of the test period for the forecast horizon of weeks 3~4

图 3. 在预测延伸期为第 3~4 周中, 两种模型在测试期间内每个月平均 CS 值(a)以及温度异常平均(b)

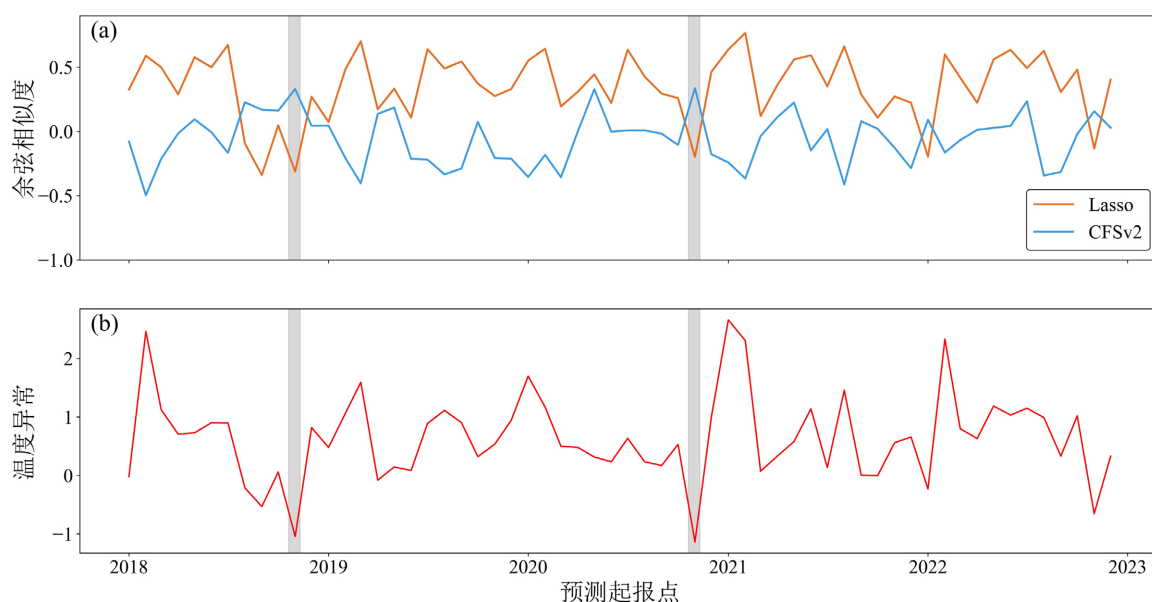


Figure 4. Average CS of two models (a), and average temperature anomalies (b) in each month of the test period for the forecast horizon of weeks 5~6

图 4. 在预测延伸期为第 5~6 周中, 两种模型在测试期间内每个月平均 CS 值(a)以及温度异常平均(b)

尽管 Lasso 模型在大多数月份的平均 CS 显著高于 CFSv2, 但是 Lasso 模型也存在某些不足。具体来说, 在温度异常值呈现出显著负异常的月份(图 3(b)和图 4(b)), CFSv2 的平均 CS 却更高。类似的结果也 He 等人指出[20]。他们发现, 在 2019 年美国大寒潮期间, SubX 多集合平均的动力模型的预测性能显著高于机器学习模型。这一现象的可能原因在于, 极端天气事件的稀有性导致数据驱动的机器学习模型往往缺乏足够的训练数据去捕捉极端气温与各预测因子之间关系。研究表明, 土壤湿度 sm 是次季节温度预测的重要可预测性来源[31]。Boenson [24]等学者的研究指出, sm 对次季节温度预测的贡献在高温条件下尤为显著。这一发现也解释了 Lasso 回归模型在极端低温条件下预测性能相对较低的原因: 由于 sm 在低温环境中的预测贡献度减弱, 导致模型在极端低温情况下的预测能力受到限制。但是, 由于动力模型遵循物理规律, 因而能够较好地模拟和预测极端天气事件的发生。

为了进一步探究 Lasso 和 CFSv2 预测性能的差异, 我们计算了 Lasso 和 CFSv2 在全国所有格点上预测结果的 CS 之间的差值。具体来说, 对于预测延伸期为第 3~4 和第 5~6 周, Lasso 分别在大约 91%和 81%的格点上表现出优于 CFSv2 的预测性能。这进一步表明了, Lasso 在中国次季节温度预测上的预测技能优于 CFSv2。

5. 结论

为了提升中国次季节温度预测能力并探索机器学习方法的应用潜力, 本研究基于经典机器学习算法 Lasso, 构建了覆盖中国所有格点的次季节温度预测模型, 并采用余弦相似度指标评估 Lasso 和 CFSv2 动力模型在 2018~2022 年测试期内的预测性能表现。结果表明, Lasso 在测试期上的整体预测性能优于 CFSv2, 其在所有格点上的平均 CS 值相比于 CFSv2 提高了 0.33 (预测延伸期为第 3~4 周)和 0.34 (预测延伸期为第 5~6 周)。Lasso 模型在全国多数格点上的预测结果 CS 值都高于 CFSv2。此外, Lasso 模型在极端低温情景下的预测技巧略低于 CFSv2, 揭示了机器学习模型在物理机制表征方面仍存在一定局限性。因此, 考虑将动力模型和机器学习模型混合建模或通过数据增强和模型融合等方法提高 Lasso 模型在极端低温情景下的预测能力。

参考文献

- [1] Rahmstorf, S. and Coumou, D. (2011) Increase of Extreme Events in a Warming World. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 17905-17909. <https://doi.org/10.1073/pnas.1101766108>
- [2] 廖承红. 全球城市升温速度惊人[J]. 生态经济, 2022, 38(1): 5-8.
- [3] Yang, Z., Wang, Q. and Liu, P. (2018) Extreme Temperature and Mortality: Evidence from China. *International Journal of Biometeorology*, **63**, 29-50. <https://doi.org/10.1007/s00484-018-1635-y>
- [4] White, C.J., Carlsen, H., Robertson, A.W., Klein, R.J.T., Lazo, J.K., Kumar, A., et al. (2017) Potential Applications of Subseasonal-to-Seasonal (S2S) Predictions. *Meteorological Applications*, **24**, 315-325. <https://doi.org/10.1002/met.1654>
- [5] 赵昌帅. 黑龙江省极端温度与降水特征及其对春玉米产量的影响[D]: [硕士学位论文]. 哈尔滨: 东北农业大学, 2023.
- [6] Barnston, A.G., Tippet, M.K., L'Heureux, M.L., Li, S. and DeWitt, D.G. (2012) Skill of Real-Time Seasonal ENSO Model Predictions during 2002-11: Is Our Capability Increasing? *Bulletin of the American Meteorological Society*, **93**, 631-651. <https://doi.org/10.1175/bams-d-11-00111.1>
- [7] Lorenc, A. (1986) Analysis Methods for Numerical Weather Prediction. *Quarterly Journal of the Royal Meteorological Society*, **112**, 1177-1194. <https://doi.org/10.1256/smsqj.47413>
- [8] Tao, L., Cui, Z., He, Y. and Yang, D. (2024) An Explainable Multiscale LSTM Model with Wavelet Transform and Layer-Wise Relevance Propagation for Daily Streamflow Forecasting. *Science of the Total Environment*, **929**, Article ID: 172465. <https://doi.org/10.1016/j.scitotenv.2024.172465>
- [9] 陶俐芝. 月降水量预报的多尺度支持向量机模型[D]: [硕士学位论文]. 长沙: 湖南师范大学, 2018.
- [10] 凌铭, 肖丽英, 赵嘉, 等. 基于 SVM-CEEMDAN-BiLSTM 模型的日降水量预测[J]. 人民珠江, 2023, 44(9): 61-68.

- [11] Merryfield, W.J., Baehr, J., Batté, L., *et al.* (2020) Current and Emerging Developments in Subseasonal to Decadal Prediction. *Bulletin of the American Meteorological Society*, **101**, E869-E896.
- [12] 孙可可, 吴小飞. FGOALS-f2 气候预测系统对西南地区持续性极端降水的次季节预测评估[J]. 高原山地气象研究, 2024, 44(1): 21-30.
- [13] He, S., Li, X., DelSole, T., Ravikumar, P. and Banerjee, A. (2021) Sub-Seasonal Climate Forecasting via Machine Learning: Challenges, Analysis, and Advances. *Proceedings of the AAAI Conference on Artificial Intelligence*, **35**, 169-177. <https://doi.org/10.1609/aaai.v35i1.16090>
- [14] He, R., Zhang, L. and Chew, A.W.Z. (2024) Data-Driven Multi-Step Prediction and Analysis of Monthly Rainfall Using Explainable Deep Learning. *Expert Systems with Applications*, **235**, Article ID: 121160. <https://doi.org/10.1016/j.eswa.2023.121160>
- [15] 成玉祥, 肖丽英, 王萍根, 等. 基于 Attention-BiLSTM 混合模型的月尺度降水量预测[J]. 人民珠江, 2024, 45(6): 73-81.
- [16] Mouatadid, S., Orenstein, P., Flaspohler, G., Cohen, J., Opreescu, M., Fraenkel, E., *et al.* (2023) Adaptive Bias Correction for Improved Subseasonal Forecasting. *Nature Communications*, **14**, Article No. 3482. <https://doi.org/10.1038/s41467-023-38874-y>
- [17] National Academies of Sciences, Engineering & Medicine (2016) Next Generation Earth System Prediction: Strategies for Subseasonal to Seasonal Forecasts. National Academies Press. <https://doi.org/10.17226/21873>
- [18] Faijaroenmongkol, T., Sarinnapakorn, K. and Vateekul, P. (2023) Sub-Seasonal Precipitation Bias-Correction in Thailand Using Attention U-Net with Seasonal and Meteorological Effects. *IEEE Access*, **11**, 135463-135475. <https://doi.org/10.1109/access.2023.3337998>
- [19] Kiefer, S.M., Lerch, S., Ludwig, P. and Pinto, J.G. (2023) Can Machine Learning Models Be a Suitable Tool for Predicting Central European Cold Winter Weather on Subseasonal to Seasonal Time Scales? *Artificial Intelligence for the Earth Systems*, **2**, e230020. <https://doi.org/10.1175/aies-d-23-0020.1>
- [20] He, S., Li, X., Trenary, L., Cash, B.A., DelSole, T. and Banerjee, A. (2022) Learning and Dynamical Models for Sub-Seasonal Climate Forecasting: Comparison and Collaboration. *Proceedings of the AAAI Conference on Artificial Intelligence*, **36**, 4495-4503. <https://doi.org/10.1609/aaai.v36i4.20372>
- [21] Zhou, X.X., Ding, Y.H. and Wang, P.X. (2010) Moisture Transport in the Asian Summer Monsoon Region and Its Relationship with Summer Precipitation in China. *Acta Meteorologica Sinica*, **24**, 31-42.
- [22] Xu, X., Du, Y., Tang, J. and Wang, Y. (2011) Variations of Temperature and Precipitation Extremes in Recent Two Decades over China. *Atmospheric Research*, **101**, 143-154. <https://doi.org/10.1016/j.atmosres.2011.02.003>
- [23] Wheeler, M.C. and Hendon, H.H. (2004) An All-Season Real-Time Multivariate MJO Index: Development of an Index for Monitoring and Prediction. *Monthly Weather Review*, **132**, 1917-1932. [https://doi.org/10.1175/1520-0493\(2004\)132<1917:aarmmi>2.0.co;2](https://doi.org/10.1175/1520-0493(2004)132<1917:aarmmi>2.0.co;2)
- [24] Zimmerman, B.G., Vimont, D.J. and Block, P.J. (2016) Utilizing the State of ENSO as a Means for Season-ahead Predictor Selection. *Water Resources Research*, **52**, 3761-3774. <https://doi.org/10.1002/2015wr017644>
- [25] Reynolds, R.W., Smith, T.M., Liu, C., Chelton, D.B., Casey, K.S. and Schlax, M.G. (2007) Daily High-Resolution-Blended Analyses for Sea Surface Temperature. *Journal of Climate*, **20**, 5473-5496. <https://doi.org/10.1175/2007jcli1824.1>
- [26] van den Dool, H.M., Saha, S. and Johansson, Å. (2000) Empirical Orthogonal Teleconnections. *Journal of Climate*, **13**, 1421-1435. [https://doi.org/10.1175/1520-0442\(2000\)013<1421:eot>2.0.co;2](https://doi.org/10.1175/1520-0442(2000)013<1421:eot>2.0.co;2)
- [27] Tibshirani, R. (1996) Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **58**, 267-288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- [28] Saha, S., Moorthi, S., Pan, H., Wu, X., Wang, J., Nadiga, S., *et al.* (2010) The NCEP Climate Forecast System Reanalysis. *Bulletin of the American Meteorological Society*, **91**, 1015-1058. <https://doi.org/10.1175/2010bams3001.1>
- [29] Kim, H., Webster, P.J., Curry, J.A. and Toma, V.E. (2012) Asian Summer Monsoon Prediction in ECMWF System 4 and NCEP Cfsv2 Retrospective Seasonal Forecasts. *Climate Dynamics*, **39**, 2975-2991. <https://doi.org/10.1007/s00382-012-1470-5>
- [30] Nowak, K., Webb, R., Cifelli, R. and Brekke, L. (2017) Sub-Seasonal Climate Forecast Rodeo. 2017 AGU Fall Meeting.
- [31] Liu, J. and Pu, Z. (2019) Does Soil Moisture Have an Influence on Near-Surface Temperature? *Journal of Geophysical Research: Atmospheres*, **124**, 6444-6466. <https://doi.org/10.1029/2018jd029750>