

Doi:10.11840/j.issn.1001-6392.2024.03.008

基于物理关联深度学习的海浪浪高预测方法

宋巍¹, 姜浩¹, 杜艳玲¹, 张立华², 陈刚³

(1. 上海海洋大学 信息学院, 上海 201306; 2. 海军大连舰艇学院 军事海洋与测绘系, 辽宁 大连 116018; 3. 国家海洋信息中心, 天津 300171)

摘 要: 精确的海浪有效波高 (简称浪高) 预测对于海上生产生活具有重要意义。针对现有海浪浪高预测模型对不同海洋要素间关联信息考虑不足, 以及长时序浪高数据本身存在非平稳性的问题, 本文设计了一种考虑物理约束与差值约束的海浪浪高时间序列预测方法。该方法基于风速与浪高之间的物理关联, 设计物理约束, 并通过提取差分信息设计差值约束, 结合现有基于深度学习的时间序列预测模型, 实现浪高预测。采用黄海和东海的6个不同站点浮标数据进行了大量实验。实验结果表明, 本文提出的方法可以利用海洋要素间的物理关联, 有效提高浪高预测精度, 并避免因不同要素间融合造成的信息干扰; 同时, 利用差值约束, 限制时间序列预测结果的变动范围。本文方法可以与不同类型的时间序列预测模型相结合, 显著提升原有模型的性能, 并在长时间序列的预测中体现出很好的鲁棒性, 为海洋要素预测中物理与数据驱动模型的有效结合提供了思路和验证。

关键词: 浪高预测; 物理约束; 差值约束; 时间序列

中图分类号: P731.33

文献标识码: A

文章编号: 1001-6932(2024)03-0371-11

Wave height prediction method based on physical correlation constrained deep learning

SONG Wei¹, JIANG Hao¹, DU Yanling¹, ZHANG Lihua², CHEN Gang³

(1. College of Information Technology, Shanghai Ocean University, Shanghai 201306, China;

2. Department of Military Oceanography and Hydrography & Cartography, Dalian Naval Academy, Dalian 116018, China;

3. National Marine Data and Information Service, Tianjin 300171, China)

Abstract: Accurate prediction of significant wave height (wave height in short) is of great significance to marine operations and coastal human life. In response to the insufficient attention given to the physical correlations between oceanic elements in existing prediction models, as well as the non-stationarity issue with wave height time series itself, this paper proposes a wave height prediction method based on a physical constraint and a difference constraint. This method designs a physical constraint loss function by taking into account the physical correlation between wind speed and wave height, as well as a difference constraint loss function by extracting the time difference information of wave height. The loss functions are then embedded into existing deep learning-based time series prediction models to achieve accurate wave height prediction. Comprehensive experiments were conducted using buoy data from six different observation stations in the Yellow Sea and East China Sea. The results show that the proposed method can effectively improve the prediction accuracy of wave height by utilizing the physical constraint between oceanic elements while avoiding information interference caused by covariate fusion, and using difference constraints to limit the variation range of time series prediction. This method can be combined with various types of time series prediction models to significantly improve the performance of the original model and demonstrate good robustness in long-term sequence prediction. The proposed method provides ideas and validation for the effective combination of physical and data-driven models in oceanic element prediction.

Keywords: wave height prediction; physical constrain; difference constrain; time series

收稿日期: 2023-07-04; 修订日期: 2023-08-16

基金项目: 国家重点研发项目 (2021YFC3101601)

作者简介: 宋巍 (1977—), 博士, 教授, 主要从事海洋大数据分析研究, 电子邮箱: wsong@shou.edu.cn

通信作者: 陈刚, 硕士, 高级工程师, 主要从事海洋信息化、深度学习研究, 电子邮箱: cg232@126.com

由于航运、海上生产和军事活动的需要,对于海洋环境的研究较为迫切。其中,海浪作为海洋环境中十分复杂且较为特殊的一种随机现象^[1],越来越受到人们的重视。一方面,海浪的运动过程在时间上具有一定的连续性,过去时刻的浪高值会影响到将来时刻的浪高值;另一方面,海浪会受到物理机制的影响,如由风起主导作用产生的海浪,其浪高在一定程度上会受到风场运动规律的影响^[2]。

近年来,对于海浪浪高的研究提出了不同的方法,大致可以分为以下三类:一是统计学方法,从观测数据出发,通过经验公式直接计算,如Wilson公式^[3]等;二是传统数值模型,如徐丽丽等^[4]基于第三代波浪模型WAM(Wave Model),利用SWAN(Simulating Waves Nearshore)和WW3(WAVE WATCH III)^[5]的嵌套计算,建立了一套适用于东海海区和中国近海的海浪数值预报系统;三是基于神经网络的机器学习模型,主要以卷积神经网络(Convolutional Neural Network, CNN)和循环神经网络(Recurrent Neural Network, RNN)为主,实现海浪要素预测。

深度学习在海浪预测领域发展迅速并得到广泛应用。国外, Makarynskyy等^[6]将人工神经网络(Artificial Neural Networks, ANN)用于浪高预测,在大西洋和爱尔兰海岸两个站点中取得令人满意的预测结果; Cornejo-Bueno等^[7]利用分组遗传算法(Grouping Genetic Algorithm, GCA)和极限学习机(Extreme Learning Machine, ELM)相结合的方式,在美国西海岸站点浪高预测中准确率和一致性方面都有较好的表现; Altunkaynak^[8]利用基因多层感知机(Geno-Multilayer Perceptron, GMP)对美国佛罗里达州奥基乔比湖中的四个站点进行浪高预测,对比GCA,取得较高预测精度。国内,陈希等^[9]利用ANN中的反向传播算法,建立了南海硇洲岛海区的海浪预报方案,该方案对风浪浪高预测达到了一定的精度; Fan等^[10]利用LSTM提出一种快速预测海浪浪高的方式,在长期预测中获得优越的性能;莫旭涛等^[11]将卷积神经网络CNN和LSTM相结合,在中国南海北部湾的浪高预测中取得较高的预测精度。

相较于单要素时间序列预测,近期研究表明通过融合多要素的方式可以有效获取更准确和更

可靠的信息,增强预测性能^[12]。Ouled等^[13]通过融合多个时间点的SAR图像,有效提高洪水预测的准确性;王欣睿等^[14]结合珠江口近海海域的风区和水深特点,总结出对应不同区域的风浪预报方法;宋巍等^[15]尝试将有效波高、风速等多要素作为输入,设计局部-全局编码,提取多要素的局部关联信息,以提高预测性能,但是并没有充分利用海洋观测要素间的物理关联。

本文以海浪浪高预测为目标,提出一种考虑物理约束与差值约束的时间序列预测方法。该方法从风速与浪高之间的物理关联出发,设计物理约束,提取非平稳浪高数据的差分信息,设计差值约束,将风速与差值作为协变量,进行协变量融合,与时间序列预测模型相结合,得出浪高预测结果。

1 原理与方法

现有时间序列预测模型对于海洋要素间物理关联性以及长时序海浪数据的非平稳性关注不足。为此,本文首先对海浪浪高数据集进行预处理,包括缺失值填充、异常值处理等,通过分析各海洋要素之间的相关性,选择相关性较高的风速数据作为物理约束要素,并将历史数据中当前时刻与上一时刻进行差值计算,得到浪高差值;然后,将风速数据和浪高差值数据作为海浪浪高数据的协变量,经过协变量融合后输入时间序列预测网络;最后,将本文设计的约束损失函数用于训练时间序列预测网络,输出海浪浪高预测结果,具体方法流程如图1所示。

1.1 协变量融合

传统海浪浪高预测模型以浪高序列作为目标变量,通过挖掘浪高波动规律来对浪高未来趋势进行预测。但是,影响海浪浪高的要素众多,仅考虑浪高单一变量的预测方法存在一定局限性。协变量作为不同于目标变量的独立变量,在一定范围内会对目标变量产生影响。利用相关性分析,得到与浪高相关性最强的要素是风速,通过引入风速作为协变量进行融合,可以有效改进浪高预测结果。现实中的海浪浪高实测数据并不是严格平稳的^[16],差分计算具有强大的确定性信息提取能力,通过对非平稳的浪高实测数据进行差分计

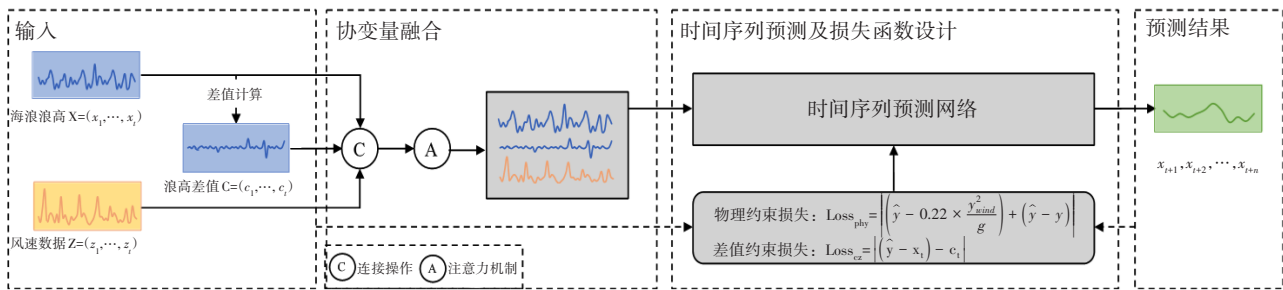


图1 方法流程图

算, 可以获取更多平稳性信息, 限制预测范围, 提高预测精度。

设海浪浪高时间序列表示为 $X = (x_1, \dots, x_t)$, 风速表示为 $Z = (z_1, \dots, z_t)$ 。计算 X 的一阶差分序列, 得到差值 $C = (c_1, \dots, c_t)$, 其中 $c_t = x_t - x_{t-1}$, $c_1 = c_2$ 以保证时间序列维度的一致性。本文将 X 定义为目标变量, Z 和 C 作为协变量, 具体融合过程如下。

首先, 通过将目标变量浪高 X 和协变量风速 Z 、差值 C 作为共同输入, 经过拼接 (Concatenate) 操作, 形成 $3 \times t$ 维向量; 其次, 考虑不同时间上对未来预测的影响, 采用注意力机制进行时间特征提取; 最后, 根据协变量对目标变量的影响程度, 设计约束损失, 以增强和利用更多有效信息来提升目标变量的预测精度。

1.2 时间注意力机制

注意力机制 (Attention Mechanism)^[17] 可以对输入序列分配不同的权重, 重点关注输入序列中的某些特定部分, 然后根据关注的区域来产生下一个输出。

具体实现方式可分为3个阶段, 第一阶段将输入映射到3个不同的空间, 得到3个向量: 查询向量 $Q \in R^{n \times d}$ 、键向量 $K \in R^{t \times d}$ 和值向量 $V \in R^{t \times d}$, 其中 n 为目标序列长度, t 为输入序列长度, d 为向量的维度; 第二阶段通过点积运算计算 K 和 Q 之间的相似系数, 使用 Softmax 激活函数归一化权重, 得到注意力权重系数: $A = \text{soft max}(QK^T)$; 第三阶段将注意力权重系数 A 与 V 加权求和, 获得最终的注意力输出, 其具体公式为:

$$\text{Attention}(Q, K, V) = \text{soft max}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (1)$$

本文采用时间注意力机制, 对输入序列的不

同时刻计算注意力权重系数 A , 来识别输入中哪些时刻对于要预测的 x_{t+1} 是更为重要的。根据在时间维度上的注意力分配, 获取不同时刻对于输出序列的重要程度, 从而更好地捕捉序列中的时间相关信息, 帮助模型更好地理解 and 处理输入数据。

1.3 基于深度学习的时间序列预测网络

时间序列预测即通过利用过去一段时间的观测数据来预测未来一段时间内的信息, 包括连续型预测 (数值预测, 范围估计) 与离散型预测 (事件预测) 等。海浪浪高在其本身的运动过程中, 有明显的时间先后顺序, 符合时间序列预测方法的研究要求。现有时间序列预测的深度学习网络种类众多, 也都取得了比较好的预测精度。为体现本文所设计方案的普适性, 选择了四类基于深度学习的时间序列预测网络作为预测模型进行验证, 分别为 RNN 类的双向长短期记忆网络 (Bi-directional Long Short-Term Memory, Bi-LSTM)^[18], CNN 类的时间卷积网络 (Temporal Convolutional Network, TCN)^[19], 参数回归类的神经元扩展分析时间序列预测模型 (Neural Basis Expansion Analysis For Interpretable Time Series Forecasting, N-BEATS)^[20], 以及编解码器类的 Transformer^[21]。

Bi-LSTM 可以更好地捕获较长时间的依赖关系, 同时捕获双向的语义依赖; 时间卷积网络 TCN 利用卷积特性, 跨时间步提取特征, 利用膨胀因果卷积 (Dilated Causal Convolutions), 可以解决长时间序列信息问题; 纯全连接的模型 N-BEATS 通过多层全连接进行时间序列分解, 每层拟合时间序列部分信息 (上一层拟合的残差), 通过分而治之的方法, 实现准确的时间序列预测; 使用自注意力机制的 Transformer 模型, 在编码器和解码器中使用自注意力层进行特征提取与复原,

解码器以自回归的方式预测未来的值,并使用注意力机制与编码器连接,以此来关注历史值中最有用的部分。

本文在现有预测模型的基础上,从物理机制出发,融合具有物理关联的协变量数据,设计约束损失函数,应用于不同时间序列预测网络。

1.4 约束损失函数

在基于深度学习的时间序列预测网络中,损失函数是用于量化预测值与真实值之间差异的非负实数函数,常见时间序列预测网络的损失函数有均方误差损失函数 $L2$,其公式如下:

$$L2 = \frac{1}{n} \sum_{i=1}^n (\hat{y} - y)^2 \quad (2)$$

式中: $\hat{y}$ 表示预测值, y 表示真实值。

该损失函数仅考虑浪高预测值与真实值之间的差异,并未充分考虑不同要素间的物理关联。本文在传统损失函数 $L2$ 的基础上融合物理约束和差值约束,以提高模型的预测精度。

1.4.1 物理约束

物理学知识通过理论和实验的方式得到明确的收集和验证,可用于抽象和总结各类自然现象和人类行为^[22],被广泛应用于许多重要的科学和工程领域中。物理约束是指将物理学的先验知识,进行高度抽象以对物体的运动和行为施加限制或限制条件。通过将物理约束与数据驱动的机器学习模型相结合,可以缓解训练数据短缺,提高预测模型的泛化性能,同时可以确保结果的合理性^[22]。

在海浪的形成过程中,由于风速不同引起水面压强变化,水面的压强造成水面的波动,从而形成不同周期和高度的海浪。根据 Phillips 等^[23]的研究,风速越大,持续时间越长,海浪就越大。由 Pierson 和 Moskowitz^[24]提出的 P-M 谱是在对充分成长的风浪记录进行谱估计和曲线的拟合时得到的,并通过北大西洋站点观测数据验证。P-M 谱是一种重力波谱,其数学表达式相对简单,而且仅与海面上方的风速有关,方便计算,因此得到广泛应用。本文使用其简化后的物理公式^[25]近似表达有效波高与风速之间的相关性,公式如下:

$$H_s \approx 0.22 \times \frac{U^2}{g} \quad (3)$$

式中: H_s 为有效波高(m); U 为 10 m 高度处的风

速(m/s); $g = 9.8$, 表示重力加速度(m/s²)。以此作为依据,结合本文数据集,设计物理约束损失 $LOSS_{phy}$ 如下:

$$LOSS_{phy} = |\hat{y} - 0.22 \times \frac{y_{wind}^2}{g}| + |\hat{y} - y| \quad (4)$$

式中: y_{wind} 表示风速实测值。

由于浪高不仅受到风速的影响,还会受到水深及海区特性等其他因素的影响,本文设计的损失函数只起约束作用,其中起主导作用的还是浪高实测数据。

1.4.2 差值约束

本文通过将当前时刻的浪高观测值与上一时刻的浪高观测值进行相减得到浪高差值 c_t ,经纬度统一后,将浪高差值作为约束变量同时输入到预测网络中。利用浪高差值修正预测结果,限制其变动范围,提高预测精度。差值约束损失 $LOSS_{cz}$ 具体计算公式如下:

$$LOSS_{cz} = |(\hat{y} - x_t) - c_t| \quad (5)$$

式中: x_t 表示预测时刻前一时刻的真实浪高值。

本文的差值约束仅考虑相邻一个时间点的真实浪高,在预测过程中仅起约束作用;而浪高预报中的误差订正是通过对更多的历史预报数据和实际观测数据进行比较,发现偏差,并对其进行修正。

1.4.3 总损失函数

基于浪高形成的物理机制构建物理约束,利用历史浪高计算浪高差值,构建差值约束。将两种不同约束损失组合得到本文约束损失,如下:

$$cLOSS = \gamma_1 LOSS_{phy} + \gamma_2 LOSS_{cz} \quad (6)$$

最终总损失函数为:

$$L = L2 + cLOSS \quad (7)$$

总损失函数采用二次方损失和一次方损失相结合的方式,考虑到二次方项对于误差较大的样本更为敏感,平方操作会放大误差的影响,而一次方项对误差较小的样本更为敏感,误差的线性关系得以保留。本文损失函数具有连续性和可导性,具有物理意义,符合实际浪高预测需求,随着模型训练的进行,损失函数的值会趋于稳定。通过相结合的方式,使得模型无论在误差较大的样本还是误差较小的样本中,都可以保持一个较好的精度。将总损失函数应用到现有的时间序列预测模型的训练中,以满足浪高预测任务的要求。

2 实验与分析

本文实验环境如下：处理器为苹果 M1 芯片；操作系统为 macOS13.0.1 (22A400)；程序设计语言为 Python3.8.1；采用 Tensorflow2.11.0 深度学习框架；集成开发环境为 PyCharm Professional 2019.3.3。损失函数中 γ_1 和 γ_2 分别在 0.01、0.1、0.2、0.5 中进行对比实验，最终 $\gamma_1 = 0.2$ ， $\gamma_2 = 0.01$ 时误差最小，于是将 γ_1 取 0.2， γ_2 取 0.01。使用 Adam (Adaptive momentum) 优化器，学习率 η 设置为 0.01。

2.1 评价指标

为了验证本文方法的有效性，选择不同类型的评价指标，包括平均绝对误差 (Mean Absolute Error, MAE)、均方根误差 (Root Mean Square Error, RMSE) 和平均绝对百分误差 (Mean Absolute Percentage Error, MAPE)。MAE、RMSE 和 MAPE 的计算公式如下所示：

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (8)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (9)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (10)$$

式中： n 为预测长度。

模型在预测浪高时，三种评价指标的数值越小，代表预测结果越准确。三种评价指标均能反映模型预测的准确度，MAPE 体现预测值与原始数据相比较的过程，可以在数量级上更直观地展示误差程度，更为公正客观。

2.2 数据集构建

本文的浪高实测数据来自国家海洋科学数据中心^[26]。数据集中的观测站点时间跨度为 2015 年

01 月 01 日 0 时至 2021 年 12 月 31 日 23 时，包含老虎滩 (LHT)、芝罘岛 (ZFD)、小麦岛 (XMD)、连云港 (LYG)、大陈 (DCH)、北碭 (BSG) 6 个海洋站的实测数据，各站点在剔除异常数据后详细情况如表 1 所示。

在本文研究中，海浪浪高、风速均给出每小时的观测值，由于浪高实测数据跨越了不同海区，其分布情况有所不同。黄海海区 LYG 和 XMD 站点的浪高值较低，且变化较为平缓，LHT 和 ZFD 站点浪高值相对较大，且呈现正偏态分布；东海海区的 DCN 和 BSG 站点浪高数据也呈正偏态分布，但受台风影响严重，浪高值变化幅度较大。

由 1.4.1 节可知，风速与浪高之间存在物理关联，计算浪高与风速之间皮尔逊相关系数， $P < 0.01$ ，二者相关系数达到 0.48，具有一定相关性。根据公式 (3) 分别对 6 个站点由风速计算浪高，并与真实浪高进行比较，得出两者分布情况近似，但计算浪高的方差偏大。以 LHT 站点为例，图 2(a) 给出浪高与风速分布情况，可以看出大部分浪高集中在 0.1~0.6 m 之间，对应风速在 1.6~5.4 m/s 之间，这也与前文物理模拟公式相对应。但是，由图 2(b) 可见，计算浪高数值范围更大。因此，本文考虑风速在预测过程中的影响时，仅将其作为损失函数的约束部分。

本文按照地理位置对不同的海区单独制作数据集，从而进行更加准确的预测。由于监测设备或者网络问题，实测数据会存在异常，对于单个时间点的异常数据，采用前后取平均的方式进行补齐；对于长时间跨度的缺失数据，使用直接截断的方式。为保证样本的时间连续性，当存在截断数据时，将截断样本直接丢弃，从后面时间连续序列重新开始。采用经典的 z-score 标准化公式 (均值为 0，标准差为 1) 对数据进行标准化。实验数据集按照 8:2 的比例划分训练集和测试集，

表 1 不同海区多源站点详细情况

海区	站点	名称	经纬度	浪高范围/m	样本数量/个
黄海	LHT	老虎滩	121.7°E, 38.9°N	0.1 ~ 3.7	51 490
	ZFD	芝罘岛	121.4°E, 37.6°N	0.1 ~ 3.6	24 877
	XMD	小麦岛	120.4°E, 36.1°N	0.0 ~ 3.3	60 131
	LYG	连云港	119.4°E, 34.8°N	0.0 ~ 2.6	45 729
东海	DCN	大陈	121.9°E, 28.5°N	0.2 ~ 7.9	46 312
	BSG	北碭	120.3°E, 26.7°N	0.1 ~ 9.1	31 053

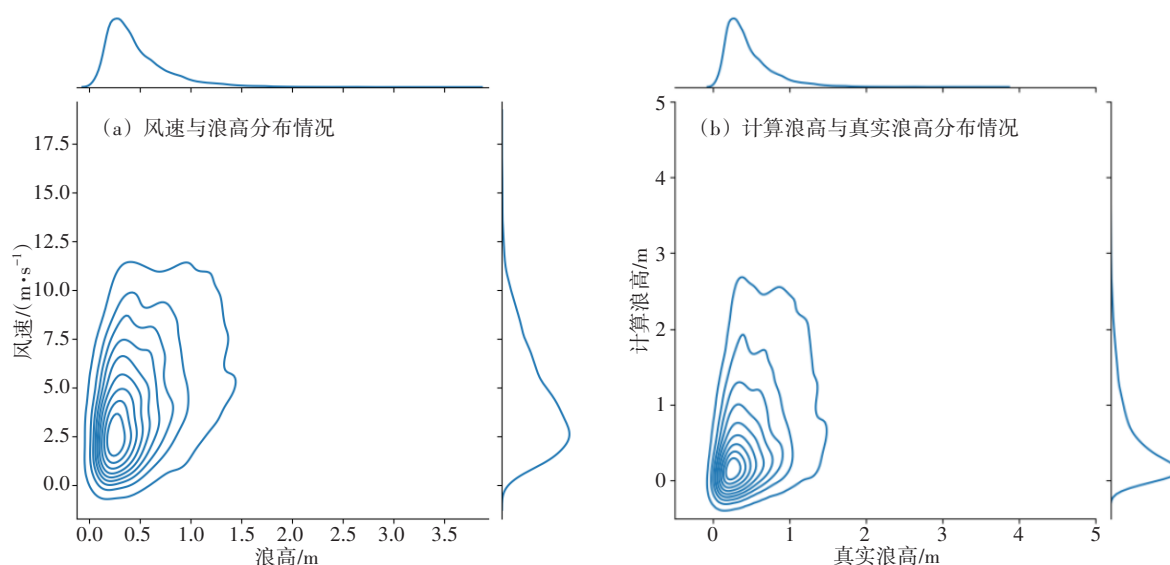


图2 LHT站点风速与浪高以及计算浪高与真实浪高分布相关性

采用滑动窗口的方式,通过对比6 h、12 h、24 h、48 h不同输入长度,选择效果最好的24 h数据作为输入,采用Sequence-Sequence的方式,分别预测未来1 h、6 h、24 h的数据。

2.3 对比实验

本文针对不同融合方式、不同预测模型以及不同预测步长单独设计对比实验,具体实验结果如下。

2.3.1 不同融合方式对比

本文将目标变量经预测网络得到出的预测结果作为对比基准方案(Plan-1),如图3所示,与不同融合方案进行横向比较。Liu等^[27]通过不同特征计算外积的方式,获得互补信息矩阵,在分类任务中都取得了较好的效果。本文基于此设计Plan-2,计算目标特征与协变量特征之间的互补信息矩阵后,与目标特征连接后输入到预测网络中,即特征层融合中的早期融合^[28],如图4(a)所示。本文还采用目标特征与协变量特征直接拼接

的方式输入到预测网络,得到预测结果,作为Plan-3,即特征层融合中的晚期融合,如图4(b)所示。

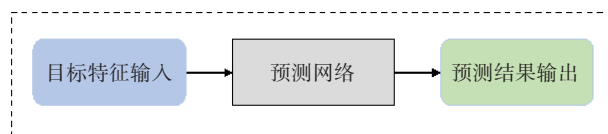


图3 基准方案 Plan-1

以黄海水区LHT和ZFD站点为例,在不同预测模型中,分别对比三种方案使用传统 $L2$ 损失函数进行网络训练后的预测结果,以及在预测网络中使用本文提出的约束损失函数($L2+cLOSS$)训练后的预测结果,实验结果评价指标如表2所示(实验结果为LHT和ZFD两站点平均结果)。

由表2可见,对比Plan-1,两种融合方式Plan-2和Plan-3使用传统 $L2$ 损失在各项指标上表现更差。这说明,通过增加与输入数据相关的协变量或交互特征,虽然理论上可以为输入数据提

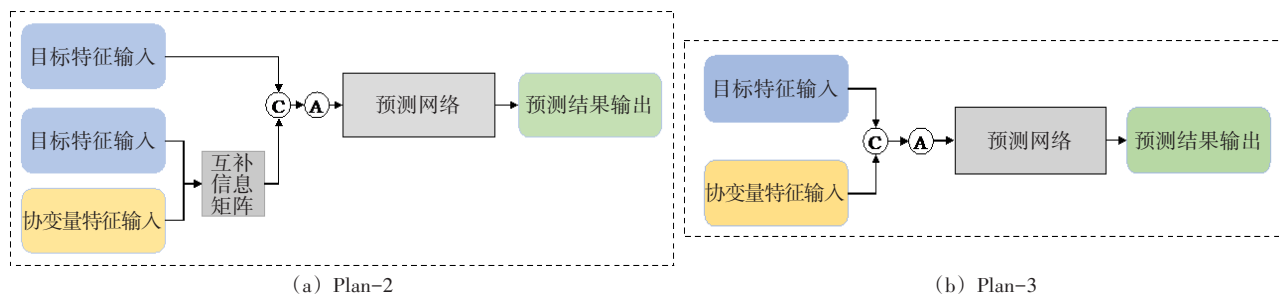


图4 不同数据融合方案

表2 不同融合方案评价指标结果

对比模型	评价指标	Plan-1 (目标变量)	Plan-2 (目标+协变量)		Plan-3 (目标+协变量)	
		<i>L2</i>	<i>L2</i>	<i>L2+cLOSS</i>	<i>L2</i>	<i>L2+cLOSS</i>
Bi-LSTM	MAE/m	<u>0.071</u>	0.081	0.077	0.073	0.064
	RMSE/m	<u>0.139</u>	0.178	0.14	0.141	0.135
	MAPE/%	<u>16.867</u>	21.224	17.527	18.256	13.518
TCN	MAE/m	0.075	0.087	<u>0.071</u>	0.081	0.067
	RMSE/m	<u>0.141</u>	0.157	0.145	0.143	0.137
	MAPE/%	<u>16.721</u>	18.937	17.825	17.825	14.385
N-BEATS	MAE/m	0.224	0.302	<u>0.199</u>	0.274	0.103
	RMSE/m	0.329	0.412	<u>0.278</u>	0.375	0.167
	MAPE/%	<u>30.955</u>	33.289	31.637	33.446	21.043
Transformer	MAE/m	0.132	0.146	<u>0.129</u>	0.138	0.114
	RMSE/m	0.204	0.198	<u>0.171</u>	0.218	0.168
	MAPE/%	31.680	29.709	<u>27.743</u>	28.985	23.414

注：黑色加粗为最佳性能，下划线为次优性能。

供更丰富的相关信息，但是简单的特征累加或交互，不仅不能给输入数据的挖掘与预测提供有用特征，还可能会对原输入数据造成不必要的干扰，严重影响模型预测的准确性。通过增加约束损失后，Plan-2和Plan-3对比*L2*损失函数训练下的结果，在各项误差评价指标上都有明显降低。这说明有效利用协变量的约束是至关重要的。

本文考虑风速在浪高的预测过程中只起约束作用，并不作为主要影响。Plan-2早期融合的方式通过构建交互信息矩阵增加了特征向量的维度，信息冗余问题严重，造成模型过大，训练困难的问题。本文设计Plan-3晚期融合的方式，输入模型前并没有产生过多的干扰信息，输入模型后再进行融合，性能上超过了单要素的Plan-1。因此，下文将在Plan-3方案的基础上使用不同预测模型进行实验，以检验该方法的泛化性。

2.3.2 不同预测模型对比

以Plan-3为基础，在不同海区多源站点中进行单步预测。Bi-LSTM、TCN、N-BEATS及Transformer四种模型的评价指标结果如表3所示。不同模型网络参数设置如下：Bi-LSTM模型设置为2层，第1层设置返回序列，神经元数量分别为128、64，激活函数使用Tanh，Dropout设置为0.2；TCN模型的核大小设置为4，过滤器数量为10，扩张卷积大小分别为1、2、4、8，激活函数使用Relu；N-BEATS模型类型设置为Generic，堆叠层邻域大小设置为24，使用共享权重；Transformer模型注意力

头数为4，Head_size为256，编码器层数为4，解码器层数为4，激活函数为Relu，Dropout设置为0.25。

从评价指标结果可以看出，所选四种不同类型的预测模型在使用本文所设计方法（融合风速数据和约束损失函数）后，对原始预测模型的性能提升显著。四种不同的预测模型中，预测性能最好的是Bi-LSTM模型，所有海区平均MAE、RMSE和MAPE指标分别降低11.203%、2.144%和19.338%；其次是TCN模型，平均MAE、RMSE和MAPE指标分别降低10.625%、1.94%和12.819%；Transformer模型表现最差，根据文献^[29]，受训练数据量的影响，Transformer模型可能并不是解决时序预测问题的较优方法。

由于本文设计的约束损失只是考虑到风速这一协变量的影响，但是浪高同时还会受到海区环境和海底地形的影响，在不同海区的预测性能差异较大。以Bi-LSTM模型为例，在黄海海区的误差普遍较大，东海海区误差较小，黄海的ZFD站点，MAPE指标达到16.35%，而在东海的BSG站点MAPE指标仅有7.27%。

进一步对比不同模型在异常和平稳海况下的表现，分别选择了东海海区BSG站点2018年7月10日06时至2018年7月12日05时（48 h）“玛莉亚”台风发生期间的数据（浪高范围1.4~9.6 m），以及黄海海区XMD站点2018年7月28日10时至2018年7月30日09时（48 h）的数据（浪高范围

表3 各模型在不同海区多源站点评价指标结果

海区	站点	评价指标	Bi-LSTM	Bi-LSTM +	TCN	TCN +	N-BEATS	N-BEATS +	Trans	Trans+
黄海	LHT	MAE/m	0.053	0.049	0.054	<u>0.051</u>	0.273	0.058	0.107	0.093
		RMSE/m	<u>0.081</u>	0.080	0.084	<u>0.081</u>	0.387	0.085	0.153	0.135
		MAPE/%	12.977	10.687	12.046	<u>10.884</u>	25.785	13.036	24.127	21.104
	ZFD	MAE/m	0.089	0.078	0.096	<u>0.082</u>	0.175	0.147	0.156	0.135
		RMSE/m	0.196	0.190	0.199	<u>0.193</u>	0.271	0.249	0.255	0.202
		MAPE/%	20.757	16.349	21.396	<u>17.885</u>	36.125	29.049	39.232	25.724
	XMD	MAE/m	0.057	0.050	0.063	0.058	0.058	<u>0.052</u>	0.062	0.074
		RMSE/m	0.092	<u>0.093</u>	0.096	0.106	0.097	0.098	0.098	0.109
		MAPE/%	13.268	11.121	15.783	12.817	14.498	<u>12.005</u>	14.633	17.163
	LYG	MAE/m	0.039	0.031	0.045	<u>0.033</u>	0.145	0.110	0.068	0.051
		RMSE/m	<u>0.076</u>	<u>0.076</u>	0.084	0.075	0.129	0.118	0.088	0.077
		MAPE/%	21.048	14.466	22.094	<u>16.828</u>	42.576	38.800	28.796	22.229
东海	DCN	MAE/m	0.103	0.091	0.118	0.112	0.109	<u>0.098</u>	0.168	0.129
		RMSE/m	0.162	0.161	0.211	0.201	0.201	<u>0.191</u>	0.216	0.193
		MAPE/%	10.172	8.591	11.794	10.993	10.279	<u>9.383</u>	20.288	12.091
	BSG	MAE/m	0.099	0.096	0.102	<u>0.098</u>	0.398	0.331	0.153	0.138
		RMSE/m	0.188	<u>0.171</u>	0.178	0.178	0.267	0.193	0.173	0.105
		MAPE/%	8.473	7.274	7.705	<u>7.594</u>	26.932	22.959	11.810	10.498

注：Trans 表示 Transformer 模型，各个模型对比基准选用 Plan-1 单变量预测结果，其中“+”表示将本文所设计 Plan-3 和约束损失函数应用到对应模型中。黑色加粗为最优性能，下划线为次优性能。

0.1~0.5 m)，进行预测，对比图如图 5 所示。总体上，Bi-LSTM/Bi-LSTM+、N-BEATS/N-BEATS+和 TCN/TCN+模型预测结果与真实浪高拟合度明显优于 Transformer/Transformer+；相较原始模型，使用本文的约束损失后（Bi-LSTM+、N-BEATS+、TCN+、Transformer+），预测值与真实浪高之间拟合效果有明显提升，这与表 3 中的评价指标一致。在存在台风的 BSG 站点中，N-BEATS/N-BEATS+模型对极值浪高的预测性能最好，经损失约束后的 TCN+模型对极值预测效果也有明显改善。在更为平稳的 XMD 站点序列中，Bi-LSTM/Bi-LSTM+模型表现更好，能够捕捉到海浪浪高的微小变化；N-BEATS/N-BEATS+和 TCN/TCN+模型的预测结果接近，但 N-BEATS+的结果相较于 TCN+更稳定。就浪高变化的整体趋势而言，Bi-LSTM+模型仍然是效果更好的。由于台风等异常海况出现属于少数情况，因此表 3 中 Bi-LSTM 模型的总体性能指标最好。

2.3.3 不同预测步长对比

根据实际需求，在海浪浪高的预测过程中，需要进行长时间序列的预测。以 Bi-LSTM 预测模型为例，将本文提出的 plan-3 结合约束损失函数

应用到 6 h 和 24 h 的多步长时间序列预测任务中，即直接预测未来 6 h 和 24 h 的浪高。对比 L2 损失函数，在不同海区的多源站点中，不同预测步长的 MAPE 评价指标如图 6 所示。

可以看出，无论在 1 h 的预测任务还是 6 h、24 h 的多步预测任务中，使用本文方法，误差评价指标都有所减小，特别是在长时间序列预测上对原预测模型的性能提升明显。如，LHT、LYG、DCN、BSG 四个站点能将 6 h、24 h 的多步预测结果提高到与 1 h 接近。而 ZFD 和 XMD 站点在 6 h、24 h 的预测中性能与 1 h 预测有差距，但其 MAPE 指标下降了 4%~6%，与其他各站点的平均值一致。这两个站点误差较高可能与站点数据本身有关，由于原模型在 6 h、24 h 预测上的误差率远高于其他站点，模型使用本文方法优化后，整体误差依然偏大。

2.4 消融实验

以 Bi-LSTM 作为预测模型进行消融实验，在 1 h 预测中，以 DCN 站点为例，分别设计了没有物理约束损失的 w/o $LOSS_{phy}$ ，没有差值约束的 w/o $LOSS_{cz}$ ，没有注意力机制的 w/o Attention，与具有完整模块的模型做对比，评价指标结果如表 4 所示。

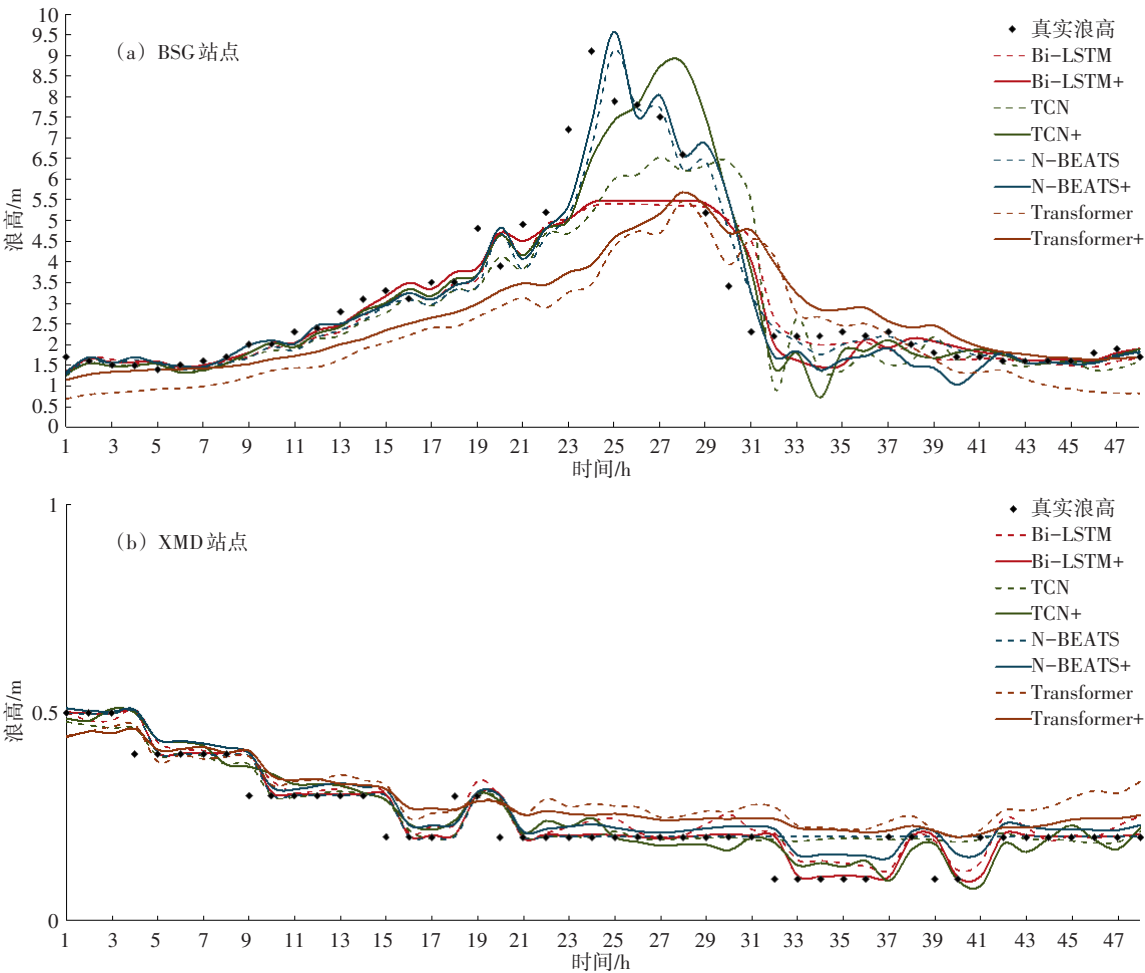


图5 不同模型单步预测48 h对比

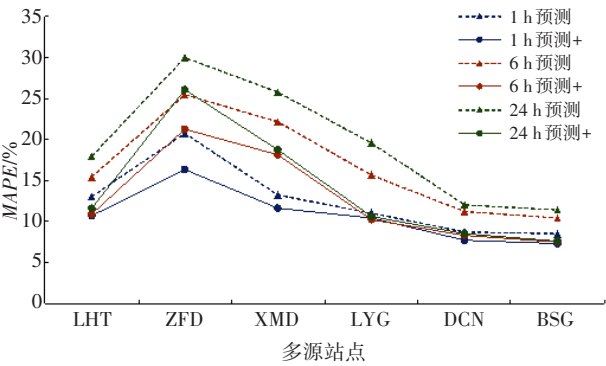


图6 不同预测步长MAPE对比图

注：“+”表示使用本文所设计Plan-3结合约束损失函数后的评价指标。

可以看出，影响最为明显的是物理约束和差值约束。通过增加物理约束，有效利用了要素间的物理关联；通过增加差值约束，限制预测结果的范围，进而提高预测精度。通过两种约束相结合的方式，能获得更加准确的预测效果。

表4 基于Bi-LSTM消融实验评价指标结果

模型	MAE/m	RMSE/m	MAPE/%
w/o $LOSS_{phy}$	0.100 26	0.185 51	9.985 56
w/o $LOSS_{CZ}$	0.098 12	0.187 75	9.409 99
w/o Attention	0.097 61	0.183 74	9.167 29
Bi-LSTM+	0.090 73	0.161 39	8.590 77

注：黑色加粗为最佳性能数据。

3 结语

海浪浪高作为一种重要的海洋要素，精确、快速的预测对于航海安全、海岸防护、海上风电场等具有重要意义。结合深度学习技术在时间序列预测领域的发展，本文依据不同海区的多源海洋站实测数据，通过增加协变量的方式，融合多要素数据，设计两种约束损失，应用到不同预测模型中，实现对浪高的有效预测。

本文主要的结论如下:

(1) 目前基于深度学习的海浪浪高预测方法未有效利用海洋要素的物理特性, 本文通过将风速作为协变量, 设计物理约束, 有效利用风速与浪高之间的物理关联; 同时考虑到海浪浪高的非平稳性, 提取差值信息并设计差值约束, 限制浪高预测结果的变动范围, 提高预测精度。本文方法可以与Bi-LSTM、TCN、NBEATS和Transformer等不同模型相结合, 对比原始网络, 均能取得更低的预测误差, 平均MAPE指标分别为11.415%、12.834%、20.872%和18.135%。

(2) 不同的融合方式会对预测精度产生影响。综合考虑浪高数据特点, 对比特征层融合中的早期融合方式, 本文设计的晚期融合方式不仅可以避免因为融合造成的要素间信息干扰, 同时能够取得更高的预测精度, 更适合于海浪浪高预测。

(3) 各海区站点海况有所不同, 模型选择对于预测结果的优劣也有所差异。在浪高变化相对平稳的站点, 如XMD、ZFD中, TCN和Bi-LSTM模型预测结果拟合度更高; 而在变化较为剧烈的站点, 如BSG、DCN中, NBEATS模型效果更好。

海浪浪高预测会受到多种客观环境的影响, 本文仅考虑了风速这一项协变量。海浪成因复杂, 不仅会受到水深及海区特性等因素的作用, 此外, 在不同季节海浪分布特征也有所差异。在未来的研究中, 将尝试加入其他要素特征, 对不同海区、不同季节进行单独分析, 增强对海浪生成复杂性的理解, 进一步提高海浪浪高预测的准确性和可靠性。

参 考 文 献

- [1] 许富祥. 海浪预报现状与未来[J]. 海洋预报, 2005(S1): 172-175.
- [2] 杨永增, 乔方利, 夏长水, 等. 海浪对海洋上层的动量与混合作用分析[J]. 海洋科学进展, 2003(4): 363-368.
- [3] WILSON E B, HILFERTY M M. The distribution of chi-square[J]. Proceedings of the National Academy of Sciences, 1931, 17(12): 684-688.
- [4] 徐丽丽, 肖文军, 石少华, 等. WaveWatch III和SWAN模型嵌套技术在业务化海浪预报系统中的应用及检验[J]. 海洋通报, 2015, 34(3): 283-294.
- [5] GROUP T W. The WAM model: a third generation ocean wave prediction model[J]. Journal of Physical Oceanography, 1988, 18(12): 1775-1810.
- [6] MAKARYNSKY O. Improving wave predictions with artificial neural networks[J]. Ocean Engineering, 2004, 31(5-6): 709-724.
- [7] CORNEJO-BUENO L, NIETO-BORGE J C, GARCÍA-DÍAZ P, et al. Significant wave height and energy flux prediction for marine energy applications: A grouping genetic algorithm - Extreme Learning Machine approach[J]. Renewable Energy, 2016, 97: 380-389.
- [8] ALTUNKAYNAK A. Prediction of significant wave height using geno-multilayer perceptron[J]. Ocean Engineering, 2013, 58: 144-153.
- [9] 陈希, 沙文钰, 李妍, 等. 人工神经网络技术在海浪预报中的应用[J]. 海洋通报, 2002(2): 11-15.
- [10] FAN S, XIAO N, DONG S. A novel model to predict significant wave height based on long short-term memory network[J]. Ocean Engineering, 2020, 205: 107298.
- [11] 莫旭涛, 李自立. 神经网络北部湾海洋浪高预测[J]. 现代电子技术, 2021, 44(13): 55-59.
- [12] 余辉, 梁镇涛, 鄢宇晨. 多来源多模态数据融合与集成研究进展[J]. 情报理论与实践, 2020, 43(11): 169.
- [13] SCHAIER M O, HAMMAMI I, FOUCHER S, et al. Flood extent mapping from time-series SAR images based on texture analysis and data fusion[J]. Remote Sensing, 2018, 10(2): 237.
- [14] 王欣睿, 娄安刚, 孙波涛. 珠江口近海海域海浪预报方法研究[J]. 海洋通报, 2009, 28(3): 113-116.
- [15] 宋巍, 赵勐, 贺琪, 等. 多要素局部-全局特征关联的有效波高预测模型[J/OL]. 上海海洋大学学报, 2023, 32(3): 449-459.
- [16] 武雨晴, 陶爱峰, 文斌, 等. 基于长期连续测波资料的波浪平稳性分析[C]. 中国海洋工程学会. 第十七届中国海洋(岸)工程学术讨论会论文集(上). 北京: 海洋出版社, 2015.
- [17] CHAUDHARI S, MITHAL V, POLATKAN G, et al. An attentive survey of attention models[J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2021, 12(5): 1-32.
- [18] ZHOU P, SHI W, TIAN J, et al. Attention-based bidirectional long short-term memory networks for relation classification[C]. Proceedings of the 54th annual meeting of the association for computational linguistics (volume 2: Short papers). 2016: 207-212.
- [19] BAI S, KOLTER J Z, KOLTUN V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling[J/OL]. arXiv e-prints. 2018; p. arXiv: 1803.01271.
- [20] ORESHKIN B N, CARPOV D, CHAPADOS N, et al. N-BEATS: Neural basis expansion analysis for interpretable time series forecasting[C]. In 8th International Conference

- on Learning Representations(ICLR), 2020.
- [21] WU N, GREEN B, BEN X, et al. Deep transformer models for time series forecasting: The influenza prevalence case[J/OL]. arXiv: 2001.08317, 2020.
- [22] MENG C, SEO S, CAO D, et al. When physics meets machine learning: A survey of physics-informed machine learning[J/OL]. arXiv preprint arXiv: 2203.16797, 2022.
- [23] PHILLIPS O M. On the generation of waves by turbulent wind[J]. Journal of Fluid Mechanics, 1957, 2(5): 417–445.
- [24] THORSOS E I. Acoustic scattering from a "Pierson – Moskowitz" sea surface[J]. The Journal of the Acoustical Society of America, 1990, 88(1): 335–349.
- [25] PIERSON JR W J, MOSKOWITZ L. A proposed spectral form for fully developed wind seas based on the similarity theory of SA Kitaigorodskii[J]. Journal of Geophysical Research, 1964, 69(24): 5181–5190.
- [26] 国家海洋科学数据中心. 波浪与风场数据[DB/OL]. [2022-07-02]. <http://mds.nmdis.org.cn>.
- [27] LIU Z, SHEN Y, LAKSHMINARASIMHAN V B, et al. Efficient low-rank multimodal fusion with modality-specific factors[C]. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL), Melbourne, Australia, July 15–20, 2018, 2247–2256.
- [28] ZADEH A, CHEN M, PORIA S, et al. Tensor fusion network for multimodal sentiment analysis[C]. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 2017, 1103–1114.
- [29] ZENG A, CHEN M, ZHANG L, et al. Are transformers effective for time series forecasting[C]. In the 37th AAAI Conference on Artificial Intelligence (AAAI-23), Washington DC, USA, 2023, 11121–11128..
- (本文编辑：袁泽轶)