基于多周期趋势分解和两级融合策略的浪高预测方法

郑小罗[1], 李其超[1], 姜 浩[1], 宋 巍[1*], 邓小东[2]

（1. 上海海洋大学 信息学院，上海 201306；
2. 国家海洋局 东海预报中心，上海 200081）

摘 要：不同海区的近岸海浪浪高具有明显差异性。针对当前大部分时间序列预测模型缺乏对不同地区（多源）浪高预测的适应性难题，提出了一种基于局部加权回归的多周期趋势分解（Seasonal and Trend decomposition using Loess，STL）和两级融合策略的浪高预测模型，简称为 MSTL-WH（Multiple STL-Wave Height）。结合多源近岸浪高时间序列的多周期性、非线性和非平稳性的特点，首先利用周期图法提取多源近岸浪高数据集中的 4 个主要周期，并基于主要周期进行多次 STL 分解，将复杂的原始浪高序列分解为周期项、趋势项和余项；然后利用长短期记忆网络（Long Short Term Memory，LSTM）并结合两级融合策略，搭建近岸浪高预测网络；最后使用自注意力机制重新调整权重并输出未来 12 h 的浪高值。通过与当前主流时间序列预测方法对比，验证了所提方法在多源近岸浪高序列预测中具有较好的实用性和更低的预测误差。

关键词：近岸浪高预测；周期趋势分解；长短期记忆网络；两级融合策略

中图分类号：P731.33 **文献标志码：**A **文章编号：**1671-6647（2023）03-0466-11

doi: 10.12362/j.issn.1671-6647.20220809001

引用格式： 郑小罗, 李其超, 姜浩, 等. 基于多周期趋势分解和两级融合策略的浪高预测方法[J]. 海洋科学进展, 2023, 41(3): 466-476. ZHENG X L, LI Q C, JIANG H, et al. Wave height prediction method based on multi-period trend decomposition and two-level fusion strategy[J]. Advances in Marine Science, 2023, 41(3): 466-476.

浪高是全球大气系统中的一个重要参数，由于海洋环境极其复杂，预测浪高是一项非常具有挑战性的任务[1]。一般地，浪高的预测方法大致可以分为 3 类：数值波浪模型、经典线性时间序列模型和机器学习模型。

数值波浪模型基于能量平衡方程，例如第三代波浪模型 WAM（Wave Model）、基于 WAM 的近岸波浪模拟 SWAN（Simulating Waves Nearshore）和 WW3（WAVE WATCH Ⅲ）[2]，其利用预报的风场资料，将第三代生成波模式应用于大尺度、长时间的浪高预测。然而，数值波浪模型是物理规律驱动的数值逼近模型，具有输入复杂度高、对风场资料的要求高及计算过程耗时长等特点，比较适用于大区域海浪的反演和预报。

由数据驱动的方法中，经典线性时间序列模型包括自回归模型（Autoregressive，AR）、自回归移动平均模型（Autoregressive Moving Average，ARMA）和差分整合移动平均自回归模型（Autoregressive Integrated Moving Average，ARIMA）等[3]。Soares 和 Ferreira[4] 将 AR 用于浪高预测；Ge 和 Kerrigan[5] 对

收稿日期：2022-08-09
资助项目：国家重点研发计划项目（2021YFC3101601）；上海市科委地方能力建设项目（20050501900）
作者简介：郑小罗（1981—），男，工程师，硕士，主要从事海洋数据分析、海洋信息化技术方面研究.
E-mail: xlzheng@shou.edu.cn
***通信作者：**宋 巍（1977—），女，教授，博士，主要从事海洋大数据分析、计算机视觉方面研究.
E-mail: wsong@shou.edu.cn

（李 燕 编辑）

比了 AR 和 ARMA 用于海浪预报的性能，在 10 种不同的海浪数据测试下，ARMA 的误差和效率均优于 AR。Agrawal 和 Deo[6] 采用 ARMA 和 ARIMA 在不同预测区间进行在线浪高预测，结果表明 ARMA 和 ARIMA 的性能非常相似。但是 ARMA 模型对序列的平稳性有要求，而 ARIMA 能够通过差分运算将非平稳数据转换成平稳数据，解决平稳性依赖问题，所以 ARIMA 模型是应用更为广泛的经典线性时间序列模型。

近年来，以大数据为基础的机器学习算法表现出强大的预测能力。Makarynskyy[7] 和 Cornego-bueno 等[8] 利用人工神经网络（Artificial Neural Networks，ANN）和极限学习机（Extreme Learning Machine，ELM）预测浪高，在准确率和一致性方面都有较好的表现。由于支持向量机（Support Vector Machine，SVM）和支持向量回归（Support Vector Regression，SVR）具有强大的泛化能力和完整的数学理论，在海浪预测方面也得到广泛应用[9-11]。Malekmohamadi 等[12] 对 SVM、ANN、贝叶斯网络和自适应神经模糊推理系统（Adaptive neural fuzzy Inference System，ANFIS）的浪高预测效果进行了深入研究，结果表明，ANN、ANFIS 和 SVM 的预测结果都在可接受范围，而贝叶斯网络的预测结果相对不可靠。陆小敏等[13] 通过集成栈式自编码器和 XGBoost（eXtreme Gradient Boosting）[14] 算法，进行有效波高预测，但是并没有考虑到数据分布的不平衡和数据集的规模大小。黄心裕等[15] 基于 Prophet 算法[16] 对海南近海波高和周期进行了分析和预测。Prophet 算法是基于时间序列分解（Seasonal、Trend 及 Residual）和机器学习的拟合实现的，适用于具有强烈季节性影响的时间序列分析，对数据异常和趋势变化具有稳健性。深度学习作为机器学习的子领域，极大地促进了计算机视觉和自然语言处理的发展。长短时记忆网络（Long Short Term Memory，LSTM）是最流行的时间序列预测深度学习网络之一，被广泛地用于预测风速、交通、太阳能和股票价格等[17] 领域。莫旭涛和李自立[18] 将卷积神经网络（Convolutional Neural Network，CNN）和 LSTM 相结合，在中国南海北部湾的浪高预测取得较高的预测精度。顾兴健等[19] 基于 LSTM 建立的海浪环境短期预报模型，在我国海域上的预测结果取得了较好的精度，但是随着时间长度的推移，误差会逐渐增大。此外，许多新的深度学习网络，例如 DeepAR[20] 也被广泛应用到时间序列预测相关领域，但是在浪高预测中应用较少。

上述方法尝试从多个角度进行浪高预测，但未能充分考虑浪高序列的周期性、非平稳性和非线性，一定程度地限制了浪高预测的精度。基于局部加权回归（Locally Weighted Regression，Loess）的周期趋势分解（Seasonal and Trend decomposition using Loess，STL）[21] 能够探索复杂的时序数据的规律，适用于时序预测任务，且具有较强的鲁棒性。为了实现对多站点未来 12 h 的浪高预测，本文针对多源浪高时序数据的多周期性、非平稳性和非线性等特点，提出了一种基于多周期 STL 分解和两级融合策略的浪高预测方法 MSTL-WH。首先通过周期图法估计多源浪高数据集中的主要周期，然后利用 STL 将复杂的浪高序列分解为简单的趋势项、周期项和余项，最后使用 LSTM 结合两级融合策略以融合长中短时序下的浪高特征并进行浪高预测。应用实测浪高数据开展实验，并与经典时间序列模型 ARIMA、加性时间序列预测模型 Prophet、集成学习模型 XGBoost、长短时记忆网络 LSTM 和基于概率预测深度学习模型 DeepAR。实验结果表明，MSTL-WH 能够对不同站点未来 12 h 的浪高进行准确预测，综合表现优于 ARIMA、Prophet、XGBoost、LSTM 和 DeepAR 方法，各级浪高下 MAE 均低于 20%，符合业务化运行标准。

1　方　法

MSTL-WH 的流程如图 1 所示。首先对浪高数据集进行预处理，包括缺失值填充、异常值处理等。使用周期图法提取浪高数据集中的 4 个主要周期，以代表不同监测站点的浪高周期，并根据这

4 个主要周期对浪高序列进行 4 次 STL 分解，将复杂的浪高序列分解为 12 个简单分量；然后采用结合两级融合策略的 LSTM 网络进行特征提取与融合，有效学习不同周期、不同时间海浪信息对未来浪高的影响；最后使用全连接层并结合注意力机制输出浪高预测值，实现对多站点浪高的精确预测。

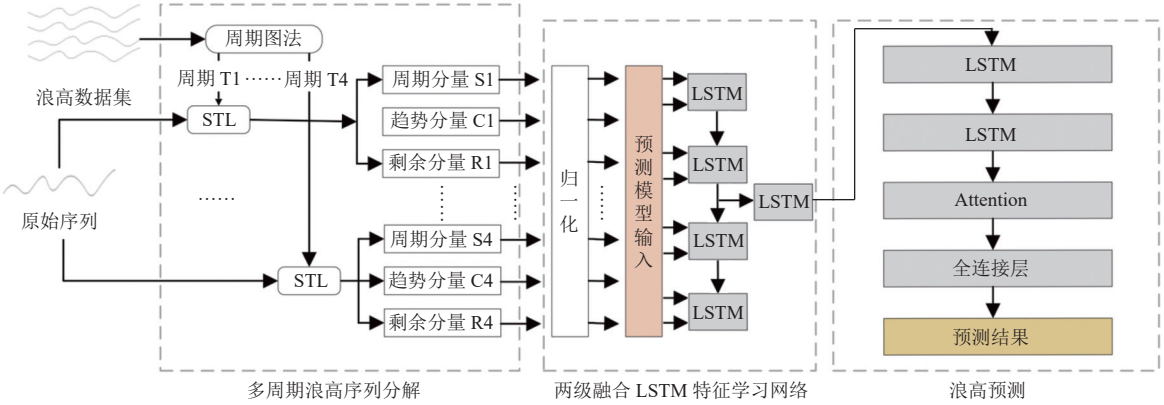

图 1 方法流程图

Fig. 1 Flow chart of the proposed method

1.1 基于 STL 的多周期-趋势分解

近岸浪高受多重因素影响，使得浪高难以预测，近岸浪高具有一定的周期性，可以使用 STL 将浪高时间序列分解为 3 个简单分量：趋势分量、周期分量和剩余分量。不同站点浪高时间序列一般具有不同的周期，因此本文利用从数据集中提取的 4 个主要周期，对每一条浪高时间序列进行 4 次 STL 分解，提取其不同周期下的分量，从而提高预测方法的泛化性能，实现对多源浪高时间序列的精准预测。

浪高数据主要周期提取的方法采用经典的周期图法，即信号功率谱密度估计方法。对于一个浪高序列$X(n)$，对其进行离散傅里叶变换为$X\left(\mathrm{e}^{\mathrm{j}\omega}\right)$，该序列的周期图可以定义为：

$$\widehat{P}(\omega)=\frac{1}{N}\left|X\left(\mathrm{e}^{\mathrm{j}\omega}\right)\right|^2。\tag{1}$$

根据式(1)得到的功率谱估值，找出最大功率对应的频率，并取其倒数得到该序列的主要周期。

STL 是由 Cleveland 等[21] 提出的一种经典的时间序列分解方法，其优点在于，对带有异常值的时间序列分解出的分量有更强的鲁棒性，本文中，STL 将浪高序列分解为 3 个分量的分解表达如下：

$$X_i=C_i+S_i+R_i \qquad (i=1,2,3,\cdots,N)\,,\tag{2}$$

式中：X_i为浪高序列；C_i为趋势分量；S_i为周期分量；R_i为剩余分量；N为序列长度。STL 是一种迭代方法，其分解过程主要分为内循环和外循环，内循环嵌套在外循环中。内循环用于更新趋势分量和周期分量，外循环用于计算稳健的权值，通过在下一次内循环中使用这些权值，以减少异常值对更新后续内循环中趋势分量和周期分量的影响。内循环是 STL 的主要部分，具体循环过程如下。

第 1 步：去趋势性。在内循环迭代$k+1$次，使用在第k次迭代时得到的趋势分量T_i^k对原始浪高序列进行去趋势：$X_i^{'}=X_i-T_i^k$。

第 2 步：平滑周期性。使用 Loess 对去趋势浪高序列$X_i^{'}$的每个周期子序列进行平滑，得到浪高序列的临时周期分量$\widetilde{S}_i^{k+1}$。

第 3 步：低通滤波处理。对第 2 步中得到的临时周期分量$\widetilde{S}_i^{k+1}$使用低通滤波器处理，之后使用 Loess 对其进行平滑得到$\widetilde{T}_i^{k+1}$。

第 4 步：临时周期分量去趋势性。将第 2 步中得到的临时周期分量$\widetilde{S}_i^{k+1}$减去$\widetilde{T}_i^{k+1}$防止低频信号进入周期分量，得到第$k+1$次循环的周期分量S_i^{k+1}，表示为$S_i^{k+1}=\widetilde{S}_i^{k+1}-\widetilde{T}_i^{k+1}$。

第 5 步：去周期性。原始浪高序列X_i减去周期分量S_i^{k+1}得到去周期序列$X_i^{''}$，表示为$X_i^{''}=X_i-S_i^{k+1}$。

第 6 步：平滑趋势性。对第 5 步得到的去周期序列$X_i^{''}$使用 Loess 平滑，得到第$k+1$次循环的浪高趋势分量T_i^{k+1}。

第 2、3、4 步是内循环的周期平滑部分，第 6 步是趋势平滑部分。在外循环中，利用内循环所得的浪高趋势分量和浪高周期分量来计算浪高剩余分量$R_i^{k+1}=X_i-S_i^{k+1}-T_i^{k+1}$。浪高序列 4 个主要周期经过 STL 分解，被分解为 12 个简单分量，接下来以这 12 个分量作为神经网络预测模型的输入。

1.2　两级融合策略

对于不同周期 STL 分解的 12 个分量，直接利用 LSTM 进行特征学习可能会破坏各周期的独立性，造成多周期之间的干扰。为此，结合两级融合的思想设计特征学习网络，网络第一层使用 4 个 LSTM 网络分别提取 4 个不同周期输入子序列的特征。

LSTM 网络是 RNN 的一种改进模型，相较于 RNN 模型，它具有特殊的记忆和遗忘模式，每一个 LSTM 网络拥有一个记忆单元（cell），在t时刻的状态记为c_t，其值通过输入门、遗忘门和输出门更新，它们一般使用 sigmoid 或者 tanh 函数进行激活。记忆单元的工作流程为：在时刻t，记忆单元通过 3 个门接收当前状态x_t与上一时刻记忆单元的隐藏状态h_{t-1}，除此之外，每一个门还接收记忆单元的状态c_{t-1}。接收到输入信息之后，每一个门对不同来源的输入进行计算，并且由其激活函数决定其是否激活。输入门的输入经非线性变换后，与经过遗忘门处理过的记忆单元状态进行组合，产生新的记忆单元状态c_t。最后，记忆单元状态c_t通过非线性运算和输出门的控制产生记忆单元的输出h_t。

从浪高时间序列的长、中、短周期考虑，结合两级融合的思想设计预测网络，网络第一层使用 4 个 LSTM 网络分别提取输入序列中 4 个子序列的特征，然后将输出融合到一个 LSTM 网络中，之后再连接 2 层 LSTM 网络进一步提取特征（LSTM 网络中的隐藏单元数量均为 128）。

1.3　自注意力机制

对浪高特征进行充分提取之后，每个特征已经获得了一定的权重，为进一步强化特征，提升模型预测精度，使用自注意力模块对权重进一步分配。

注意力机制本质为一个查询对应多个键值对的映射。注意力机制的原理可分为 3 个阶段：第 1 阶段计算每一个查询 $\boldsymbol{Q}$ 和各个键 $\boldsymbol{K}$ 的相似度，获得每个键对应值 $\boldsymbol{V}$ 的权重；第 2 阶段使用 Softmax 激活函数归一化权重，从而获得权重系数；第 3 阶段对权重系数和对应的键值 V 进行加权求和，获得最终注意力权值，其计算式为：

$$\mathrm{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\mathrm{Softmax}\left(\frac{QK^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V},\tag{3}$$

式中：$\boldsymbol{Q}\in\mathbb{R}^{n\times d_k}$；$\boldsymbol{K}\in\mathbb{R}^{n\times d_k}$；$\boldsymbol{V}\in\mathbb{R}^{n\times d_k}$，$n$为输入序列长度，$d_k$为输入序列特征维度。

自注意力机制也被称作内部注意力机制，是一种特殊的注意力机制。一般情况下，注意力机制中的输入和输出是不一样的，比如在翻译任务下，输入序列是需要翻译的英文语句，输出序列则是翻译后的中文语句，注意力机制发生在输出的 $\boldsymbol{Q}$ 和输入的所有元素之间。对于自注意力机制则是发生在输入序列或输出序列的内部元素之间，也就是式（3）中 $\boldsymbol{Q}$、$\boldsymbol{K}$ 和 $\boldsymbol{V}$ 三个矩阵都来源于相同的

输入序列。自注意力机制可以减少对外部信息的依赖，将输入序列不同位置的信息关联，擅长提取数据的内部相关性。

2 实验验证

实验验证了 MSTL-WH 在测试集和新数据集上的预测误差，并通过与 ARIMA、XGBoost、Prophet 和 LSTM 这 4 个具有代表性的预测方法对比，证明本文所提方法的优越性。本文实验硬件环境如下：处理器为英特尔酷睿 i9-10885H，CPU 频率为 2.4 GHz，内存为 16 GB；操作系统为 Windows10（64 位）；程序设计语言为 Python3.6.9（64 位），集成开发环境为 PyCharm Professional 2019.3.3；程序设计过程中的 LSTM 网络由 Keras2.1.5 程序包实现。整个浪高预测实验主要分为 4 部分：①数据集构建；②消融实验；③对比实验；④泛化性能验证。

2.1 评价指标

为验证 MSTL-WH 的有效性，选择平均绝对百分误差（Mean Absolute Percentage Error，MAPE）和平均绝对误差（Mean Absolute Error，MAE）作为评价指标，对比不同预测模型预测浪高的性能。MAPE 和 MAE 的公式如下所示：

$$\mathrm{MAPE} = \sqrt{\frac{\sum_{i=1}^{N_t}\left(\tilde{y}_i - y_i\right)^2}{N_t}}, \tag{4}$$

$$\mathrm{MAE} = \frac{100}{N_t}\sum_{i=1}^{N_t}\frac{\left|y_i - \tilde{y}_i\right|}{y_i}, \tag{5}$$

式中：y_i为真实浪高；$\tilde{y}_i$为预测浪高；N_t为预测长度。模型在预测浪高时，MAPE 和 MAE 的数值越小，代表预测结果越准确。

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[22]。由于浪高监测设备异常或网络异常等原因，实测数据会包含异常值，因此需要对实测数据进行异常值处理，具体处理方式为：对于异常值连续长度不超过 3 的浪高序列，使用异常点前后的 3 个有效值的均值进行填充；对于异常值连续长度超过 3 的浪高序列进行截断保存。对所有站点的实测数据进行上述处理后，得到 94 条长度不一的浪高序列，这些序列用于后续数据集的制作。

数据集中的观测站点分布在 24°～39°N，包含小长山（XCS）、小麦岛（XMD）、连云港（LYG）、大陈（DCH）、老虎滩（LHT）、芝罘岛（ZFD）、南麂（NJD）七个海洋站的实测数据，每条浪高数据时间跨度为 2017 年 1 月至 2020 年 4 月，采样频率为 1 次/h，共 29 160 h。

由于浪高实测数据跨越了不同海区，其浪高数据的周期性变化也大不相同。为了确定对浪高数据进行 STL 分解的周期值，使用周期图法对每一条浪高序列进行功率谱密度估计。提取功率最大的频率，取其倒数作为候选周期，共获得 94 个候选周期，各观测站点浪高时序数据的主要周期如图 2 所示。结合候选周期的直方图，最终确定 140、221、314 和 460 h 作为周期值。根据上述 4 个周期，STL 将每一条复杂的浪高序列分解成 12 个简单分量，将每个分量归一化至区间（0，1）后用于后续的数据集制作。为了捕捉不同站点的海浪周期，从而更加准确地预测浪高，在制作数据集时，滑动窗口的长度应大于海浪序列数据的周期。因此，在使用滑动窗口切分浪高序列时，窗口大小设置为 500 h，步长为 2 h，前 488 个时间点的浪高值及其分量作为输入，后 12 个时间点的浪高值作为真值，对所有序列进行滑动窗口切分后，共得到 69 836 个样本。按照 8∶2 划分训练集和测试集，得到

55 868 个样本的训练集、13 967 个样本的测试集，数据形状分别为(55 868，488，12)和(13 967，12)，训练集中各等级浪高样本的数量如图 3 所示。

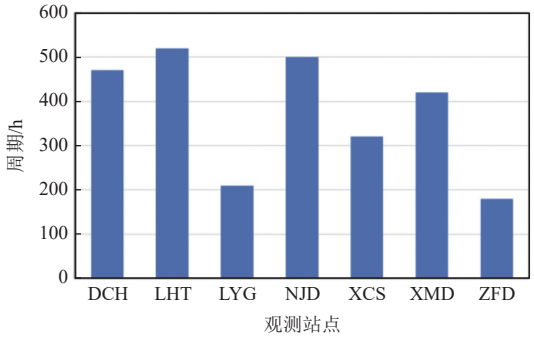

图 2　各观测站点浪高时间序列数据的主要周期

Fig. 2　Periods of the wave height time series at different observation stations

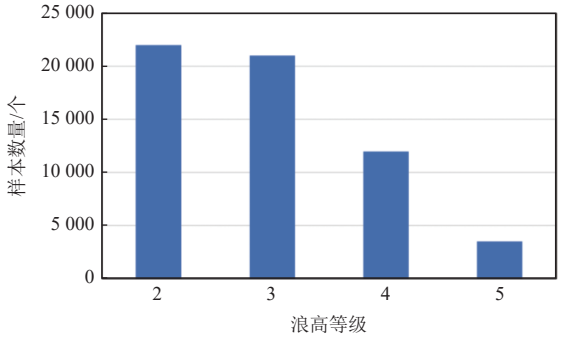

图 3　训练集中各等级浪高的样本数量

Fig. 3　Number of samples at different levels of wave height in the training data set

2.3　消融实验

首先通过对比实验，确定一个具有 3 层 LSTM 的预测网络，基于此网络对比采用单周期 STL 分解前后和多周期 STL 分解的浪高预测误差，验证多周期 STL 分解的有效性，然后分别在是否采用多周期 STL 分解的基础上，对比采取两级融合前后的浪高预测误差，证明两级融合的有效性，最后通过实验确定融合层后 LSTM 的层数。

为了验证多周期 STL 分解的有效性，向一个 3 层 LSTM 预测网络分别输入原始浪高时间序列、单周期 STL 分解后的浪高时间序列和多周期 STL 分解后的浪高时间序列，对比浪高预测误差，其中每层 LSTM 的隐藏神经元个数为 128。未采用 STL 分解时，模型输入数据为原始浪高序列，维度为(488，1)；采用单周期 STL 分解时，使用周期图法得到数据集中功率最大周期，并据此对输入的浪高序列进行 STL 分解，模型的输入数据维度为(488，3)；采用多周期 STL 分解时，模型的输入数据维度为(488，12)。模型的输出均为未来 12 h 的浪高值，结果如表 1 所示。

表 1　多周期 STL 分解对预测性能的影响

Table 1　Impact of multi-period STL decomposition on prediction performance

前处理方式	MAE/m	MAPE/%
未经前处理	0.19	40.25
单周期 STL 分解	0.17	36.85
多周期 STL 分解	<u>0.15</u>	<u>30.97</u>

注：下划线标注数据为最佳性能数据。

从表 1 可以看到，使用单周期进行 STL 分解在一定程度上可以减小预测误差，相比之下，多周期进行 STL 分解可以更进一步减小预测误差，使 MAE 降至 0.15 m，MAPE 降至 30.97%，这表明多周期 STL 分解对于减小预测误差发挥了重要作用。

接下来为了验证两级融合的有效性并进一步说明多周期 STL 分解的有效性，我们分别对单周期 STL 分解及多周期 STL 分解下是否采用两级融合对预测误差的影响进行了对比。在融合层之后连接 1 层 LSTM 网络和 1 层全连接层来输出浪高预测值，其中每层 LSTM 的隐藏神经元个数为 128。实验结果如表 2 所示。

从表 2 可以看出，无论在单周期 STL 分解下还是在多周期 STL 分解下，采用两级融合均可明显减小预测误差。相对单周期 STL 分解，采用多周期 STL 分解具有更小的预测损失，结合两级融合可将 MAE 降低至 0.11 m，MAPE 降低至 20.56%。因此，多周期 STL 分解及两级融合对于多源浪高时间序列预测任务具有明显提升效果。

表 2 单/多周期 STL 分解下两级 LSTM 特征融合对预测性能的影响

Table 2 Impact of two-stage LSTM feature fusion on prediction performance under single/multi-period STL

前处理方式	是否两级融合	MAE/m	MAPE/%
单周期 STL 分解	否	0.17	36.85
	是	0.14	26.58
多周期 STL 分解	否	0.15	30.97
	是	0.11	20.56

注：下划线标注数据为最佳性能数据。

最后，在采用多周期 STL 分解及两级融合的基础上，我们在特征融合层后连接不同层数的 LSTM，然后使用自注意力层对权重再分配，最后连接 1 个全连接层输出浪高预测值，其中每层 LSTM 的隐藏神经元个数为 128，结果如表 3 所示。

表 3 LSTM 层数对预测性能的影响

Table 3 Impact of the number of layers of LSTM on prediction performance

参 数	融合后 LSTM 层数			
	1	2	3	4
MAE/m	0.11	0.08	0.07	0.08
MAPE/%	20.56	14.05	11.87	12.72

注：下划线标注数据为最佳性能数据。

由表 3 可以看出，随着网络中 LSTM 层数的增加，网络的预测性能也随之提升，但当 LSTM 层数增加至 4 层时，由于网络参数量过多导致过拟合，网络性能开始下降，当 LSTM 的添加层数为 3 时，模型的预测误差最小，MAE 低至 0.07 m，MAPE 低至 11.87%，因此，在 MSTL-WH 中，使用 3 层 LSTM 进一步提取融合后的特征。

2.4 对比实验

基于时间序列的预测模型主要包括经典线性时间序列模型、机器学习模型和深度学习模型，本文从中选取了 ARIMA、Prophet、XGBoost、LSTM 和 DeepAR 这 5 个具有代表性的预测模型与 MSTL-WH 进行对比。

根据最新发布的海浪等级国家标准[23]，将模型预测结果计算平均值后分为 4 个等级区间，即 2 级对应浪高 0.10～0.50 m、3 级对应浪高 0.50～1.25 m、4 级对应浪高 1.25～2.50 m、5 级对应浪高 2.50～4.00 m，评价指标结果如表 4 所示。同时，为直观观察各模型在不同等级浪高下的预测性能，在不同浪高等级范围随机选择了 4 个时间序列，使用各种对比模型进行了 24 h 预测，预测结果如图 4 所示。

表 4 各模型在不同浪高等级下的评价指标结果

Table 4 Evaluation metrics of different models at different levels of wave height

模 型	平均 MAE/m	平均 MAPE/%	MAE/m				MAPE/%			
			2 级	3 级	4 级	5 级	2 级	3 级	4 级	5 级
LSTM	0.42	35.65	0.15	0.23	0.35	0.96	53.08	35.9	22.36	31.24
ARIMA	0.28	22.23	0.08	0.16	0.27	0.59	31.71	20.24	17.04	19.94
Prophet	0.48	46.06	0.20	0.31	0.51	0.91	81.62	39.68	32.08	30.86
XGBoost	0.21	20.15	0.09	0.14	0.2	0.42	33.29	19.28	13.08	14.95
DeepAR	0.27	19.14	0.06	0.20	0.25	0.60	16.83	25.26	17.24	17.21
MSTL-WH	0.17	12.13	0.03	0.10	0.17	0.39	15.14	13.18	8.16	12.03

注：下划线标注数据为最佳性能数据。

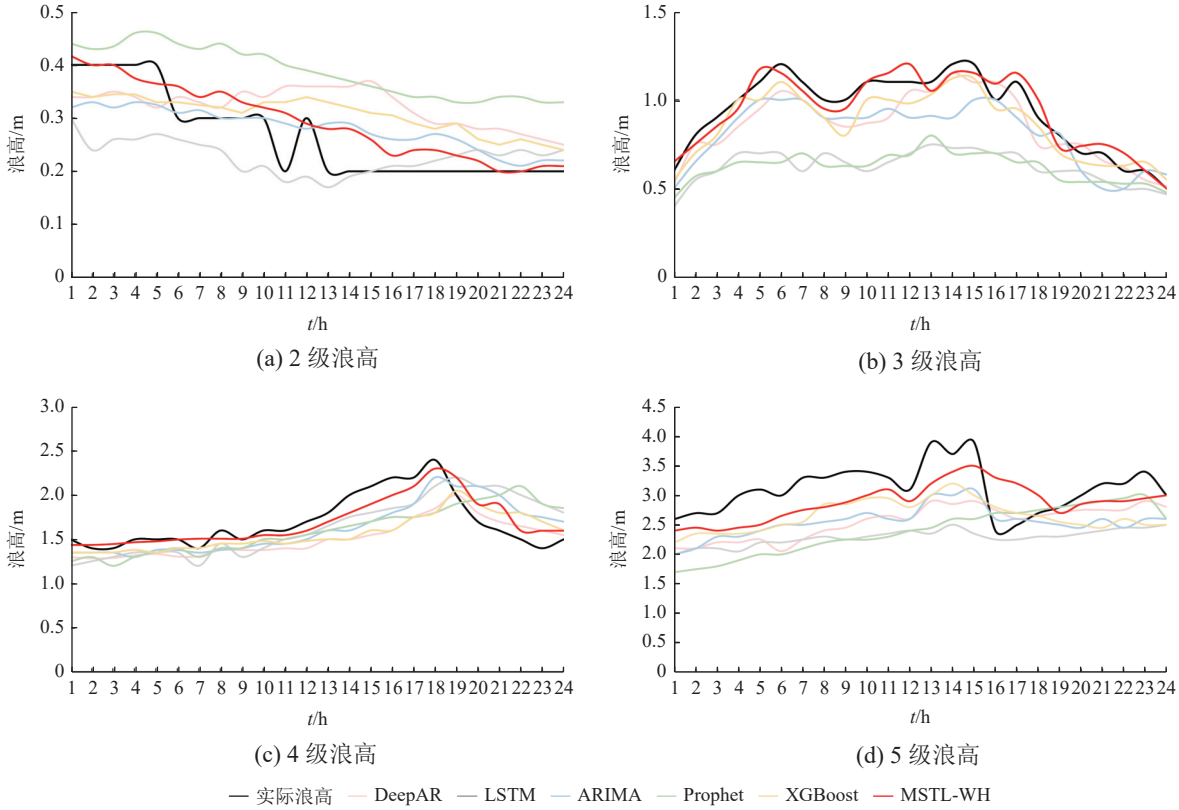

图 4　各模型不同等级浪高下的预测效果

Fig. 4　Results of different models at different levels of wave height

根据 24 h 的浪高值预测结果对比(图 4)，MSTL-WH(红色)与实际浪高值(黑色)的拟合度较高。集成学习方法 XGBoost(黄色)和经典的 ARIMA 模型(蓝色)整体上比其他 3 种方法性能更好。DeepAR 模型(粉色)整体性能与 ARIMA 模型类似。Prophet 模型(绿色)对低等级的浪高预测值偏高，但是高等级的浪高预测值偏低，说明其对数据变化适应性较差。同样，LSTM 网络(灰色)预测网络效果也不佳，总体预测值偏低，这对海浪预警非常不利。最后，大部分模型在海况 4 级的浪高范围内预测趋势较好，这可能与这一等级的训练数据较充分有关，但也存在波动峰值滞后的情况，本文模型滞后性较小。

表 4 从误差指标上进一步给出客观评价。MSTL-WH 在各等级浪高下的预测指标均优于其他模型，尤其对于 3、4 级浪高，即在 0.50～1.25 m、1.25～2.50 m 区间的浪高序列，MSTL-WH 的平均绝对误差仅为 0.10 m 和 0.17 m，在平均绝对误差±0.20 m 的误差范围内，浪高等级不会发生变化。对于 2 级浪高，MSTL-WH 的平均绝对误差仅为 0.03 m，几乎可以忽略不计。因此，本文提出的 MSTL-WH 在各等级浪高下有着良好的预测性能，满足浪高预报业务化的要求(MAPE≤20%)。

2.5　泛化性能验证

尽管 MSTL-WH 在验证集上的准确性和稳定性表现良好，为了进一步验证其泛化能力，本文在不同站点以及新数据集上分别进行了实验。各模型针对不同站点(多源)的 MAE 评价指标结果如表 5 所示。

由表 5 可以看出，MSTL-WH 在 7 个不同站点中有 5 站 MAE 评价指标上优于其他模型。在 DCH 站点，MSTL-WH 和 XGBoost 的预测指标相同，均优于其他模型。在 XMD 站点中，DeepAR 的

MAE 评价指标更小，预测性能相对更好，但 MSTL-WH 仍然是所有预测模型中排名第二的模型。总体来看，MSTL-WH 在不同站点（多源）的预测误差变化小，模型性能稳定。

表 5 各模型在不同站点（多源）的 MAE 评价指标结果（m）

Table 5 MAE evaluation metrics of different models at different stations (m)

模 型	站 点						
	DCH	LHT	LYG	NJD	XCS	XMD	ZFD
LSTM	0.32	0.16	0.11	0.49	0.14	0.16	0.12
ARIMA	0.13	0.28	0.11	1.42	0.08	0.05	0.16
Prophet	0.26	0.18	0.19	0.28	0.18	0.21	0.23
XGBoost	0.07	0.09	0.04	0.17	0.09	0.04	0.09
DeepAR	0.10	0.34	0.16	1.98	0.22	0.02	0.13
MSTL-WH	0.07	0.05	0.03	0.09	0.04	0.05	0.05

注：下划线标注数据为最佳性能数据。

对国家海洋科学数据中心与训练集一致的 7 个站点 2020 年 4 月至 5 月的浪高实测数据进行预处理（详见 2.2 节描述）后，共得到 3 620 条浪高序列，覆盖了 2～5 级的浪高。MSTL-WH 在这一新数据集上的各项指标如表 6 所示。

表 6 MSTL-WH 在新数据集上的各项指标

Table 6 Results of MSTL-WH in new dataset

参 数	浪高等级				平均值
	2 级	3 级	4 级	5 级	
MAE/m	0.05	0.11	0.11	0.53	0.19
MAPE/%	19.61	14.54	6.46	12.84	13.36

由表 6 可见，MSTL-WH 在新数据集上各等级浪高的平均相对误差均小于 20%，对比表 4 中的性能指标，MSTL-WH 除了在 2 级浪高预测有较大误差外，其余等级浪高情况下最大增加了 1.36% 的平均绝对误差。虽然 2 级浪高预测的 MAPE 为 19.16%，但其平均绝对误差仅为 0.05 m，对实际预报业务几乎没有影响。由此可见，MSTL-WH 在不同等级浪高预测上具有良好的泛化性。

3 结 语

本文基于浪高实测数据，针对多源浪高时序序列难以预测问题，提出了基于多周期 STL 分解和两级融合策略预测方法。首先使用周期图法从多源浪高时序数据集中提取主要周期，并据此对浪高序列进行 STL 分解，然后使用 LSTM 从短、中和长期不同的视野提取浪高特征，采用两级融合策略进行特征融合，最后使用全连接层输出未来 12 h 的浪高值。

本文所提出的浪高预测方法 MSTL-WH 对 2、3、4 和 5 级浪高序列预测的 MAPE 均低于 20%，满足浪高预报业务化运行的要求（MAPE≤20%）。相较于 LSTM、ARIMA、Prophet、XGBoost 和 DeepAR，MSTL-WH 能更好地预测具有不同周期的浪高序列，同时解决了对不同地区（多源）预测适应性差的问题，降低了预测模型的滞后性，且在新的浪高数据集下，具有良好的泛化性。

在充分分析时间序列数据特征的前提下，本文提出的方法不仅适用于浪高预测，还可以用于其他领域的时间序列预测任务，具有广阔的应用前景。

参考文献 (References)：

[1] ALI M, PRASAD R. Significant wave height forecasting via an extreme learning machine model integrated with im-

proved complete ensemble empirical mode decomposition[J]. Renewable and Sustainable Energy Reviews, 2019, 104: 281-295.

[2] GROUP T W. The WAM model: a third generation ocean wave prediction model[J]. Journal of Physical Oceanography, 1988, 18(12): 1775-1810.

[3] BOLLERSLEV T. Generalized autoregressive conditional heteroskedasticity[J]. Journal of Econometrics, 1986, 31(3): 307-327.

[4] SOARES C G, FERREIRA A M. Representation of non-stationary time series of significant wave height with autoregressive models[J]. Probabilistic Engineering Mechanics, 1996, 11(3): 139-148.

[5] GE M, KERRIGAN E C. Short-term ocean wave forecasting using an autoregressive moving average model[C]//The 11th International Conference on Control (CONTROL). Belfast, UK, IEEE, 2016: 1-6.

[6] AGRAWAL J D, DEO M C. On-line wave prediction[J]. Marine Structures, 2002, 15(1): 57-74.

[7] MAKARYNSKYY O. Improving wave predictions with artificial neural networks[J]. Ocean Engineering, 2004, 31(5/6): 709-724.

[8] 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.

[9] MAHJOOBI J, MOSABBEB E A. Prediction of significant wave height using regressive support vector machines[J]. Ocean Engineering, 2009, 36(5): 339-347.

[10] ELBISY M S. Sea wave parameters prediction by support vector machine using a genetic algorithm[J]. Journal of Coastal Research, 2015, 31(4): 892-899.

[11] SALCEDO-SANZ S, BORGE J C N, CARRO-CALVO L, et al. Significant wave height estimation using SVR algorithms and shadowing information from simulated and real measured X-band radar images of the sea surface[J]. Ocean Engineering, 2015, 101: 244-253.

[12] MALEKMOHAMADI I, BAZARGAN-LARI M R, KERACHIAN R, et al. Evaluating the efficacy of SVMs, BNs, ANNs and ANFIS in wave height prediction[J]. Ocean Engineering, 2011, 38(2/3): 487-497.

[13] 陆小敏, 刘凡, 蔡丽华, 等. 集成栈式自编码器与 XGBoost 的深度学习海浪有效波高预报模型[J]. 计算机与现代化, 2021(4): 32-36. LU X M, LIU F, CAI L H, et al. Ensemble stacked autoencoders and XGBoost based deep learning model for significant wave height forecasting[J]. Computer and Modernization, 2021(4): 32-36.

[14] CHEN T, GUESTRIN C. XGBoost: a scalable tree boosting system[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, 2016: 785-794.

[15] 黄心裕, 唐军, 王晓宇. 基于Prophet算法的海南近海波浪长时段时序分析与预测[J]. 海洋学报, 2022, 44(4): 114-121. HUANG X Y, TANG J, WANG X Y. Long-term time-series analysis and forecasting of Hainan offshore waves based on Prophet algorithm[J]. Haiyang Xuebao, 2022, 44(4): 114-121.

[16] TAYLOR S J, LETHAM B. Forecasting at scale[J]. The American Statistician, 2018, 72(1): 37-45.

[17] SHAHID F, ZAMEER A, MUNEEB M. A novel genetic LSTM model for wind power forecast[J]. Energy, 2021, 223: 223-241.

[18] 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. MO X T, LI Z L. Prediction of ocean wave height in Beibu Gulf by neural network[J]. Modern Electronic Technology, 2021, 44(13): 55-59.

[19] 顾兴健, 赵璐, 金明, 等. 基于 LSTM 神经网络的我国典型试航海域环境短期预报方法研究[J]. 中国造船, 2017, 58(4): 100-107. GU X J, ZHAO L, JIN M, et al. Research on the short-term forecast method of my country's typical trial sea area environment based on LSTM neural network[J]. China Shipbuilding, 2017, 58(4): 100-107.

[20] SALINAS D, FLUNKERT V, GASTHAUS J, et al. DeepAR: probabilistic forecasting with autoregressive recurrent networks[J]. International Journal of Forecasting, 2020, 36(3): 1181-1191.

[21] CLEVELAND R B, CLEVELAND W S, MCRAE J E, et al. STL: a seasonal-trend decomposition[J]. Journal of Of-

ficial Statistics, 1990, 6(1): 3-73.

[22] 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. National Marine Data Center. Wave and wind data[DB/OL]. [2022-07-02]. http://mds.nmdis.org.cn.

[23] 中国国家标准化管理委员会. 海浪等级: GB/T 42176—2022[S/OL]. (2022-12-30)[2023-01-05]. https://std.samr.gov.cn/gb/search/gbDetailed?id=F159DFC2A7D647EFE05397BE0A0AF334. China National Standardization Management Committee. The grade of wave height: GB/T 42176—2022[S/OL]. (2022-12-30) [2023-01-05]. https://std.samr.gov.cn/gb/search/gbDetailed?id=F159DFC2A7D647EFE05397BE0A0AF334.

Wave Height Prediction Method Based on Multi-Period Trend Decomposition and Two-Level Fusion Strategy

ZHENG Xiaoluo[1], LI Qichao[1], JIANG Hao[1], SONG Wei[1], DENG Xiaodong[2]

(1. *College of Information Technology Shanghai Ocean University*, Shanghai 201306, China;
2. *East China Sea Marine Forecasting Center of State Oceanic Administration*, Shanghai 200081, China)

Abstract: Nearshore wave height exhibits significant difference in different sea areas, and most of the time-series prediction models lack the adaptability to wave height prediction in different regions (i.e., multi-source wave height prediction). To solve this problem, we propose a wave height prediction model based on multi-period trend decomposition (STL) with locally weighted regression and two-level fusion strategy (MSTL-WH). Combining the multi-periodicity, nonlinearity and non-stationarity of nearshore wave height time series, periodogram method is used to extract four major periods of the wave height, and multiple STL decomposition was performed based on the major periods, decomposing the original wave height series into seasonal, trend and residual parts. Then, we used the Long Short-Term Memory Network (LSTM) combined with two-level fusion strategy build a nearshore wave height prediction network. Finally, self-attention mechanism was used to readjust the weights and the wave height was calculated for the next 12 hours. Compared with the existing time-series forecasting methods, the method we proposed has better practicability and smaller errors in nearshore wave height forecasting.

Key words: nearshore wave height prediction; period-trend decomposition; Long Short-Term Memory Network; two-level fusion strategy

Received: August 9, 2022