

刘懿纬,王魏,张淑雨,等.基于贝叶斯与元学习的氨氮浓度预测模型优化[J].华中农业大学学报,2023,42(4):236-243.  
DOI:10.13300/j.cnki.hnlkxb.2023.04.027

## 基于贝叶斯与元学习的氨氮浓度预测模型优化

刘懿纬<sup>1</sup>,王魏<sup>1</sup>,张淑雨<sup>1</sup>,孙俊洋<sup>1</sup>,李双双<sup>2</sup>

1.大连海洋大学信息工程学院,大连 116023; 2.大连鑫玉龙海洋生物种业科技股份有限公司,大连 116000

**摘要** 针对小样本数据下氨氮浓度模型预测精度不高、收敛速度较慢的问题,采用长短期记忆网络(long short-term memory, LSTM)建立氨氮浓度预测模型,并利用贝叶斯优化算法和元学习机制对模型进行优化。其中贝叶斯优化算法用来优化预测模型的超参数,同时给出模型参数的初始值,再使用Meta-LSTM算法学习模型梯度并允许优化器之间进行参数共享和更新,最终实现对氨氮浓度预测模型的优化。将该方法与LSTM、GRU和RNN模型进行对比试验,结果显示,研究所建模型对氨氮浓度预测的均方根误差、平均绝对误差和均方误差分别为0.027 6、0.023 9和0.000 76,均优于其他预测模型。表明基于贝叶斯和元学习的氨氮浓度预测模型对小样本数据建模有效,可以实现网络快速收敛,精度满足水产养殖中氨氮浓度预测需求。

**关键词** 贝叶斯优化; 氨氮浓度预测; Meta-LSTM; 元学习

**中图分类号** S959 **文献标识码** A **文章编号** 1000-2421(2023)04-0236-08

水产养殖是渔业发展中必不可少的部分,渔业的快速发展对养殖技术的要求不断提高<sup>[1-2]</sup>。在水产养殖过程中,氨氮浓度是衡量养殖水质的重要指标。如果养殖水体中氨氮含量过高,不仅对养殖对象的生存状态造成影响,还可能对周边的生态环境造成严重破坏<sup>[3-4]</sup>。因此,实现养殖水体中氨氮浓度的有效预测和控制是水产养殖管理的核心工作。

国内外学者在水质预测领域已有很多研究成果。如王魏等<sup>[5]</sup>采用随机配置网络模型对养殖水体氨氮浓度进行软测量建模,得到氨氮浓度的非线性模型;Zhou等<sup>[6]</sup>使用水质监测传感器结合鱼群算法建立了三维传感器网来对氨氮浓度进行预测;Dehghan-Shoar等<sup>[7]</sup>利用野外光谱技术,通过混合模式下不同物质与氨氮发生显色反应来分析氨氮的浓度。这些方法很好地解决了传统测量氨氮浓度(如纳氏试剂法)存在的测量过程复杂、成本高等问题。然而,这些模型由于本身的局限性,很难实现快速高效的逼近性能。如王魏等<sup>[5]</sup>提出的模型会随着神经元个数增加容易出现过拟合现象;Zhou等<sup>[6]</sup>的模型参数复杂,如果设置不当,容易偏离最优解致使该方法使用受限;Dehghan-Shoar等<sup>[7]</sup>提出的方法只适用于干扰因素少的环境,由于海水养殖环境复杂,影响显

色反应的物质很多,所以不适宜用在海水养殖中的水质监测。

近年来,随着信息技术的发展,数据可以实时获取,部分学者从数据预处理和优化模型参数入手提高模型精度。如李先鹏等<sup>[8]</sup>采用改进后的小波降噪对采集数据进行处理,并用循环神经网络提取数据特征,通过LSTM神经网络氨氮浓度模型实现实时预测;冯良记等<sup>[9]</sup>将改进遗传算法用于高非线性水质模型参数估计中,所得结果与给定的实测值吻合良好;唐亦舜等<sup>[10]</sup>通过优化调整数据段与隐含层神经元等模型参数,构建非线性自回归神经网络模型,提高预测精度。由于养殖过程中相关传感器较少,数字化程度较低,导致数据量不足,上述方法在小样本数据下无法保证模型的预测精度,同时水质受多种因子的影响,氨氮浓度的变化是复杂的非线性过程,通过数据预处理以及模型参数估计无法保证模型的泛化能力,因此,本研究提出基于贝叶斯和元学习的氨氮浓度预测模型,通过优化模型超参数降低人为调参带来的误差,在给定模型参数初始值后优化模型参数,允许优化器之间参数共享和更新,实现在小样本数据下的氨氮浓度预测模型优化,以期氨氮浓度预测提供新途径。

收稿日期:2022-09-30

基金项目:设施渔业教育部重点实验室(大连海洋大学)开放课题(202314);辽宁省教育厅青年科技人才“育苗”项目(QL201912)

刘懿纬, E-mail: liuyiwei563@163.com

通信作者: 王魏, E-mail: ww\_wangwei@dlou.edu.cn

# 1 材料与方法

## 1.1 数据来源

本试验中使用的数据来源于2017—2019年大连某养殖企业,每天监测1组,表1中为部分数据。

数据集包含日期、温度、盐度、亚硝酸盐质量浓度、氨氮质量浓度。基于前期文献<sup>[11]</sup>分析的结果,本研究将温度、盐度、pH作为输入变量,氨氮浓度作为输出变量进行建模,同时使用贝叶斯优化算法和元学习机制优化模型超参数及模型参数。在贝叶斯优化模型超参数过程中,将数据集分为两部分,其中训练集占80%,验证集占20%。在元学习机制优化模型参数过程中,将数据集划分为元训练集(Meta-train)和元测试集(Meta-test),元训练集包括 $D_{train}$ 和 $D_{test}$ ;元测试集包括 $D_{train}$ 和 $D_{test}$ ,其中元训练集占80%,元测试集占20%。

## 1.2 基于贝叶斯与元学习的氨氮质量浓度预测模型构建

1) LSTM神经网络的工作原理。氨氮浓度预测

表1 2017—2019年某车间水质数据

Table 1 Water quality data of a workshop from 2017 to 2019

日期 Data	温度/°C Temperature	盐度/% Salinity	pH	亚硝酸盐质 量浓度/ (mg/L) Nitrite concentra- tion	氨氮质量浓 度/(mg/L) Ammonia nitrogen concentra- tion
2017/08/11	26.00	31.00	7.80	0.017 1	0.099 7
2017/08/12	27.00	32.00	7.80	0.009 5	0.113 3
2017/08/13	26.60	30.00	7.80	0.013 7	0.074 5
.	.	.	.	.	.
.	.	.	.	.	.
2019/12/29	13.20	34.00	7.40	0.047 8	0.079 3
2019/12/30	13.20	35.00	7.50	0.084 3	0.065 9
2019/12/31	13.20	36.00	7.40	0.091 4	0.036 5

采用LSTM进行建模,相比传统循环神经网络,LSTM神经网络增加了记忆单元,并且引入3个门控单元实现记忆的更替。其原理图如图1所示。

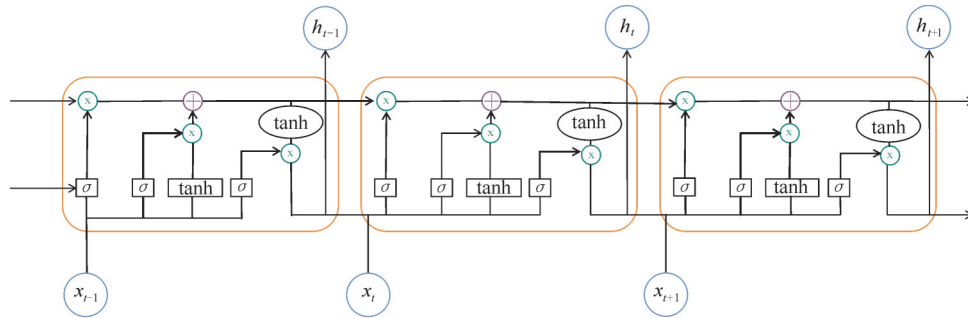


图1 LSTM神经网络原理图

Fig.1 Schematic diagram of LSTM neural network

LSTM神经网络由3个门控单元组成:

遗忘门:对记忆单元中的信息进行筛选:

$$f_t = \sigma(W_f \times [h_{t-1}, x_t] + b_f) \quad (1)$$

式(1)中, $f_t$ 是遗忘门的激活值; $W_f$ 为遗忘门的权重; $b_f$ 为遗忘门的偏置; $h_{t-1}$ 为上一时刻的隐状态; $x_t$ 为当前时刻的输入; $\sigma(\cdot)$ 为sigmoid函数。

输入门:对每个神经元的状态进行更新:

$$i_t = \sigma(W_i \times [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c \times [h_{t-1}, x_t] + b_c) \quad (3)$$

式(2)~(3)中, $i_t$ 是输入门激活值; $W_i$ 和 $W_c$ 为输入门的权重, $b_i$ 和 $b_c$ 为输入门的偏置; $C_t$ 为当前时刻的候选输出状态值。

记忆更新是通过输入门和遗忘门一起完成,新记忆主要涉及遗忘门当中留下的记忆以及输入门引入的记忆:

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (4)$$

式(4)中, $C_t$ 为当前时刻的隐藏层状态值。

输出门:对神经元当前的状态进行更新,以确定下一个神经元的记忆:

$$o_t = \sigma(W_o \times [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \times \tanh(C_t) \quad (6)$$

式(5)~(6)中, $o_t$ 是输出门激活值; $W_o$ 为输出门的权重; $b_o$ 为输出门的偏置; $h_t$ 为当前时刻的隐状态。

2) 基于贝叶斯和元学习的LSTM模型。为了实现对养殖水质数据的预测,划分数据为训练集、测试集及验证集,训练集数据的作用是对LSTM神经网络进行训练,通过L2正则化方法来避免网络出现过拟合的问题,验证集主要作为LSTM神经网络超参数的优化基础,测试集主要是针对训练之后的

LSTM网络进行精度的预测。该模型主要包括3个阶段,首先是进行LSTM模型的训练,其次是在贝叶斯优化模型中将优化需求的超参数传入以完成超参数的优化,同时完成优化的超参数对LSTM神经

网络进行训练以获得初始化参数,然后在LSTM优化器中传入,通过LSTM元学习器针对性地优化参数,最后对完成优化的模型进行精度预测,图2为基于贝叶斯和元学习的LSTM模型流程图。

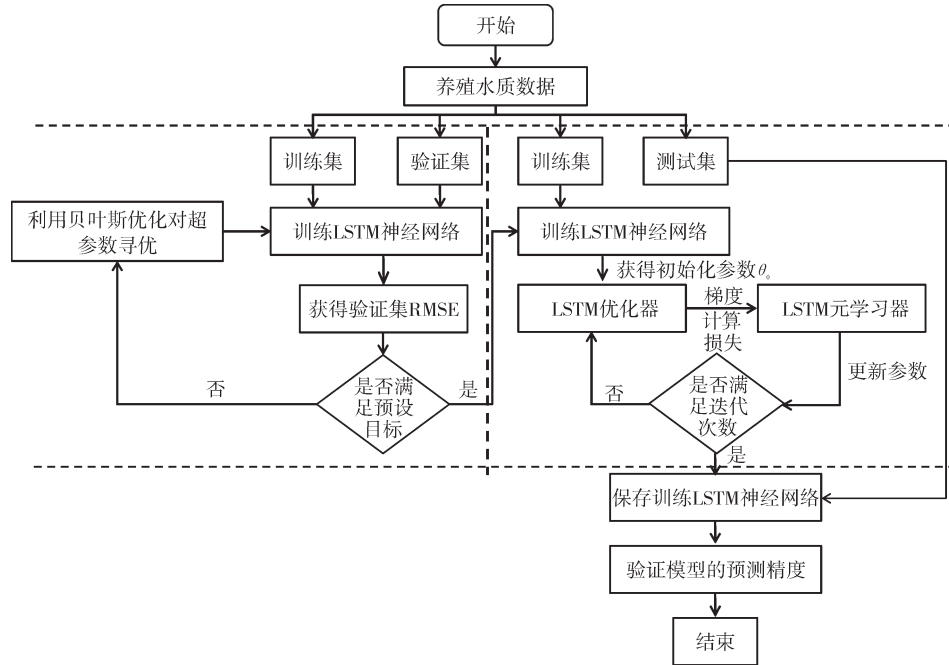


图2 基于贝叶斯和元学习的LSTM模型

Fig. 2 LSTM model based on Bayesian and Meta-learning

3) 贝叶斯优化算法。LSTM神经网络中,有一些是人为设置的超参数,比如学习率以及隐含层层数等,这些参数会对神经网络中的预测性能以及拓扑结构产生直接影响。由于调整LSTM模型参数时无法直接确定相应的目标函数,因此部分传统优化算法不适用。贝叶斯优化(BO)算法属于黑盒子优化算法,不要求目标函数的明确表达式,适用于LSTM模型的微调。

基于已知观测点和目标函数先验概率分布来进行后验概率分布的更新,再以此来寻找下一个极小值点,以确保极小值能够持续减小,保证超参数最优。贝叶斯优化目标定义为:

$$x_{\min} = \arg \min_{x \in X} f(x) \quad (7)$$

式(7)中,  $x_{\min}$  为超参数的最终优化结果,  $f(x)$  为待优化目标函数。

设待优化超参数为  $X = \{x_1, x_2, \dots, x_n\}$ , 贝叶斯优化迭代后数据集为  $D_t = \{(x_1, f(x_1)), (x_2, f(x_2)) \dots (x_n, f(x_n))\}$ 。假设高斯过程观测点服从高斯分布,表达式为:

$$f(x_{1:n}) \sim GP(\mu(x_{1:n}), \Sigma(x_{1:n}, x_{1:n})) \quad (8)$$

式(8)中,  $\Sigma(x_{1:n}, x_{1:n})$  为协方差矩阵:

$$\Sigma(x_{1:n}, x_{1:n}) = \begin{bmatrix} k(x_1, x_1) & \dots & k(x_1, x_n) \\ \vdots & & \vdots \\ k(x_n, x_1) & \dots & k(x_n, x_n) \end{bmatrix} \quad (9)$$

由贝叶斯定理可知:

$$P(f(x_{n+1}) | f(x_{1:n})) \propto P(f(x_{1:n}) | f(x_{n+1})) P(f(x_{n+1})) \quad (10)$$

持续迭代更新使  $x_{\min} = x_{t+1}$ , 最终确保超参数最优。算法流程如下:

输入: 目标函数  $f$ , 采集函数  $a$ ;

输出: 超参数向量  $x^*$ ;

① 初始化超参数向量  $x_1$ ;

② for  $t=1, 2, \dots, T$  do

③ 最大化采集函数, 得到下一个评估点:  $x_{t+1} = \arg x \in X \max a(x|D)$ ;

④ 评估目标函数值  $y_{t+1} = f(x_{t+1}) + \epsilon_{t+1}$ ;

⑤ 整合数据:  $D_{t+1} = DU(x_{t+1}, y_{t+1})$ , 同时更新概率代理模型;

⑥ end for

4) 元学习机制。元学习训练过程中, 用训练集 ( $D_{\text{Meta-train}}$ ) 训练元学习器, 使得 LSTM 神经网络优化器在测试集 ( $D_{\text{Meta-test}}$ ) 中得到较好的效果。元学习的

数据集中包含多个训练集和测试集,训练集用于优化模型参数,测试集用于对模型参数和泛化能力进行评估。

元学习在训练集中进行训练,更新后的网络参数为:

$$\theta_t = \theta_{t-1} - \alpha_t \nabla_{\theta_{t-1}} L_t \quad (11)$$

式(11)中: $\theta_t$ 是更新迭代第 $t$ 次后的参数; $\alpha_t$ 是第 $t$ 步的学习率; $\nabla_{\theta_{t-1}}$ 是更新迭代第 $(t-1)$ 次后参数 $\theta_{t-1}$ 的相应梯度, $L_t$ 是第 $t$ 次更新的损失函数。

元学习的更新和LSTM单元状态的更新相似,具体公式如(4)所示。如果令 $f_i=1, c_{t-1}=\theta_{t-1}, i_t=\alpha_t$ 以及 $c_t=\nabla_{\theta_{t-1}} L_t$ ,就与元学习的更新公式相同。

因此,用LSTM作为元学习器进行训练网络中的更新规则的学习。把LSTM中的单元状态视为元学习器参数( $C_t=\theta_t$ ),候选的单元状态是: $c_t=\nabla_{\theta_{t-1}} L_t$ 。定义了输入门 $i_t$ 和遗忘门 $f_t$ ,使元学习器基于各种更新进行学习率以及遗忘门最优参数值的确定,经过学习初始值,元学习器能够完成对最佳初始权重的确定,加快模型优化,对于输入门 $i_t$ ,它的作用相当于学习率 $\alpha_t$ ;对于遗忘门 $f_t$ ,它代表上一轮参数所占的权重,元学习器参数的状态为局部最优时,需要收缩当前参数并且遗忘之前的参数,才能更快地远离这个局部最优解,这样不仅学习了新的优化过程,还能对最优参数作为氨氮浓度模型的初始化参数进行学习,该初始化参数充当训练的最佳起点,能使优

化过程加快。具体公式如式(12)、(13)所示。

$$i_t = \sigma(W_i[\nabla_{\theta_{t-1}} L_t, L_t, \theta_{t-1}, i_{t-1}] + b_i) \quad (12)$$

$$f_t = \sigma(W_f \cdot \nabla_{\theta_{t-1}} L_t, L_t, \theta_{t-1}, f_{t-1}) + b_f \quad (13)$$

式(12)~(13)中: $i_t$ 为学习率; $\nabla_{\theta_{t-1}} L_t$ 为当前时刻的梯度, $L_t$ 为当前时刻的损失; $\theta_{t-1}$ 为上一时刻参数, $i_{t-1}$ 为上一时刻学习率; $b_i$ 为输入门的偏置; $b_f$ 为遗忘门的偏置。

5)基于贝叶斯和元学习的氨氮浓度预测模型。模型流程图如图3所示。图3中, $X_t$ 是第 $t$ 天输入变量,包括温度、pH、盐度; $Y_t$ 是输出变量,具体是氨氮浓度; $C_t$ 与 $H_t$ 分别代表运算第 $t$ 次后LSTM记忆单元中的信息以及当前运算下的隐状态; $\theta_0$ 为利用贝叶斯算法优化超参数后所得到的参数初始值; $\theta_t$ 为运算 $t$ 次之后的参数,蓝色方块表示LSTM元学习器,绿色方块表示LSTM神经网络优化器,在每次LSTM元学习器训练中,通过每时刻各个LSTM单元的状态来更新LSTM神经网络优化器的参数,LSTM神经网络优化器并未使用梯度下降,但LSTM元学习器通过梯度下降来更新参数。图3中红色虚线左边是 $D_{train}$ ,右边是 $D_{test}$ ;训练目标是在测试集 $D_{test}$ 上生成的优化器的损失 $L_{test}$ ,当在训练集 $D_{train}$ 上进行迭代时,Meta-learner在每一个 $t$ 时刻从优化器接收 $(\nabla_{\theta_{t-1}} L_t, L_t)$ 并提供新的参数集 $\theta_t$ ,这个过程重复 $t$ 次,然后在 $D_{test}$ 上评估优化并获取最终参数。

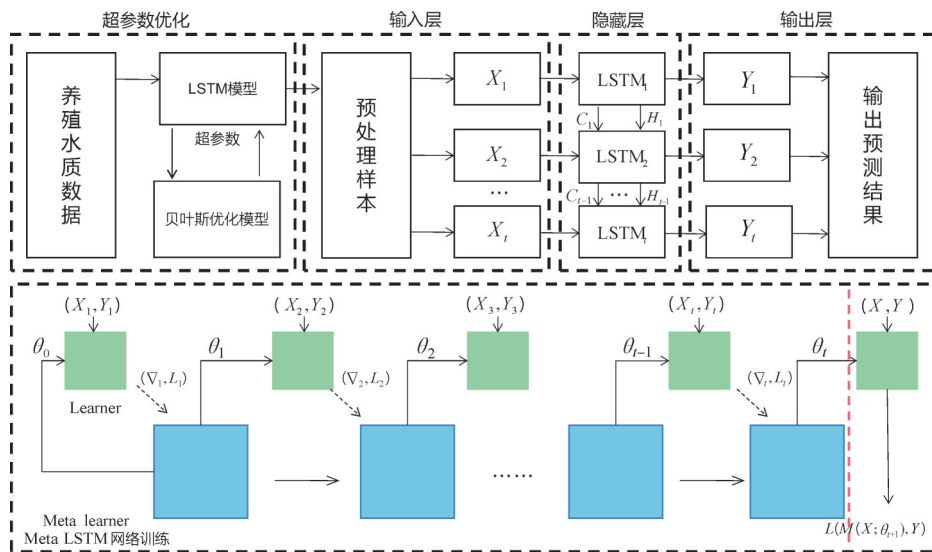


图3 基于贝叶斯和元学习的氨氮浓度预测模型流程图

Fig.3 Flow chart of ammonia nitrogen concentration prediction model based on Bayesian and Meta-learning

整体模型流程图能够划分成超参数优化、输入层、隐藏层、输出层以及Meta-LSTM网络训练5个部分。在超参数优化部分使用贝叶斯优化算法寻找

超参数,数据在输入层完成预处理,也就是划分数据集以及归一化。输出层进行输出的反归一化,计算模型的均方误差和平均绝对误差。

### 1.3 模型训练

1)模型环境搭建及试验。基于贝叶斯和元学习的预测模型环境搭建采用机器学习中的SciPy、NumPy、Matplotlib、Pandas、Scikit-learn库和深度学习中的Theano、TensorFlow、Keras库。其中LSTM元学习器(Meta-learner)进行学习率以及遗忘门值的计算,并在更新参数的值返回,每次迭代计算测试集损失,在对LSTM元学习器进行训练时,基于贝叶斯优化的LSTM模型中的权重初始化LSTM元学习器,遗忘门的偏置设为1,用于启动梯度更新,同时让LSTM元学习器使用在贝叶斯优化后所得到的学习率进行初始训练。第1组试验采用Adam-LSTM算法、SGD算法与Meta-LSTM算法进行对比,同时将LSTM模型、基于贝叶斯优化的LSTM模型(LSTM-BO)以及基于元学习的LSTM模型(LSTM-MAML)与本研究所提出的模型进行预测结果对比;第2组试验采用RNN、GRU和本研究所提出的模型进行对比。第3组试验将支持向量机(SVM)、灰色预测模型(GM)与本研究所提出的模型进行对比。建立模型之后,使用TensorFlow中的函数保存训练完成的模型参数,方便直接使用。

2)数据归一化与参数共享。本研究采用的是多输入模型,为了避免量纲所带来的影响,使取值范围不断缩小,防止过大数值给结果带来影响,对数据做归一化,具体公式为:

$$X_{std} = \frac{X - X.\min(\text{axis}=0)}{X.\max(\text{axis}=0) - X.\min(\text{axis}=0)} \quad (14)$$

$$X_{scaled} = X_{std} \times (\max - \min) + \min \quad (15)$$

式(14)~(15)中, $X.\min(\text{axis}=0)$ 是每列当中最小值构成的行向量; $X.\max(\text{axis}=0)$ 是每列中最大值所构成的行向量; $\min$ 和 $\max$ 在区间当中映射的分别是最小值以及最大值,默认值是0和1。 $X_{std}$ 为标准化结果; $X_{scaled}$ 为归一化结果。

数据归一化处理分为两步:第一标准化,对每列数据与最小值的间距以及该列数据最小值与最大值间距比例进行计算;第二在给定区间当中映射标准化数据。

为了避免元学习器(Meta-learner)的参数爆炸,本研究的模型使用一种参数在基学习器(Task-learner)的梯度坐标上共享,每个梯度坐标对应当前时刻的隐状态和单元状态值,这样LSTM元学习器学出来的更新规则使得每个维度的参数共享一套更新规则。由于共享参数,而每个维度的大小不一致,需要

对此进行预处理,将维度大小进行标准化处理来调整梯度和损失的比例,公式如下:

$$x \rightarrow \begin{cases} (\frac{\log(|x|)}{p}, \text{sgn}(x)|x| \geq e^{-p}) \\ (-1, e^p x) \end{cases} \quad (16)$$

## 2 结果与分析

### 2.1 模型参数优化

超参数优化以验证集的均方误差作为目标函数。超参数范围由经验及实验测得,隐藏层神经元数为2~20个,迭代次数为10~150次,学习率为0.0001~0.5,网络层数为1~5层。贝叶斯优化算法初始点设置为20个,迭代次数设置10次;同时构建网格搜索法(grid search, GS)和随机搜索法(random search, RS)进行对比,每次优化取最优值作为结果。考虑运行时间,对比4次试验不同方法验证集目标函数值,结果如表2所示。

表2 GS、RS与BO优化结果对比  
Table 2 Comparison of GS, RS and BO optimization results

试验次数 Times	GS	RS	BO
1	0.027	0.025	0.025
2	0.036	0.032	0.030
3	0.041	0.039	0.037
4	0.056	0.055	0.051

从表2可知,3种优化方法结果差异不大,验证集均方误差的差异均在0.01内,4次优化对比中,BO结果4次优于GS结果,1次与RS结果持平,3次优于RS结果,相较于BO方法,GS方法耗费时间,容易陷入局部最优,而RS方法与GS方法相似。

根据以上优化结果,确定本研究的模型参数为:Batch\_size=16,学习率为0.001,epoch为30,并将优化器的L2系数设为0.01。

### 2.2 模型预测结果

第1次试验采用温度、盐度、pH作为输入变量建立氨氮浓度预测模型,为了验证少量数据下Meta-LSTM算法优化效果,与Adam-LSTM、SGD算法进行对比,Meta-LSTM算法优化后的氨氮浓度预测模型在小样本预测中有一定优势,结果如图4所示。

同时将本研究所提模型(LSTM-BO-MAML)与LSTM模型、基于贝叶斯优化的LSTM模型(LSTM-BO)以及基于元学习的LSTM模型(LSTM-MAML)预测结果进行对比,本研究所提出的模型在优化超参数以及模型参数之后,预测精度

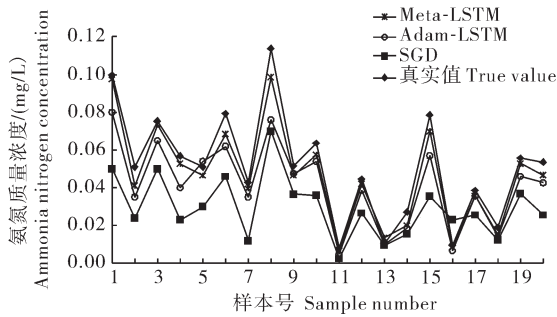


图4 以氨氮质量浓度为输出的各优化算法预测结果

Fig.4 Prediction results of each optimization algorithm with ammonia as the output

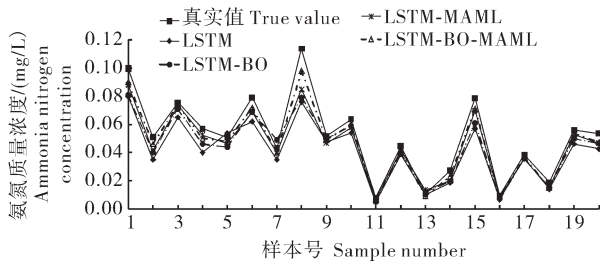


图5 以LSTM为基础的各改进模型预测结果

Fig.5 Prediction results of various improved models based on LSTM

有所提高,结果如图5所示。

### 2.3 对比试验

为了评估模型的预测效果,第2次试验将本研究提出的模型(LSTM-BO-MAML)与LSTM、RNN、GRU模型预测结果进行对比(图6),结果显示,本研究的模型预测精度较高。

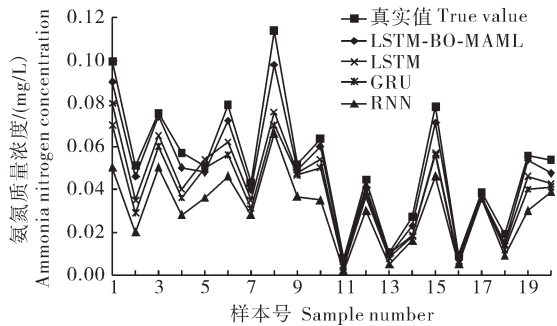


图6 常用深度学习模型预测结果

Fig.6 Prediction results of commonly used deep network models

试验中加入全部样本数据对模型进行训练后,LSTM神经网络模型的损失函数如图7所示。经过30次训练后,损失函数逐渐平稳。

为了进一步验证本研究所建模型(LSTM-BO-MAML)在小样本数据下的预测精度,与支持向量机

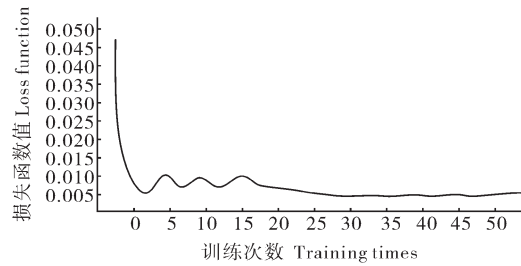


图7 损失函数图

Fig.7 Loss function diagram

模型(SVM)、灰色预测模型(GM)进行对比,将温度、盐度、pH作为输入变量,氨氮质量浓度为输出变量,LSTM-BO-MAML模型预测结果较SVM和GM模型更准确,表明本研究所建模型在小样本数据下的预测精度较高,结果如图8所示。

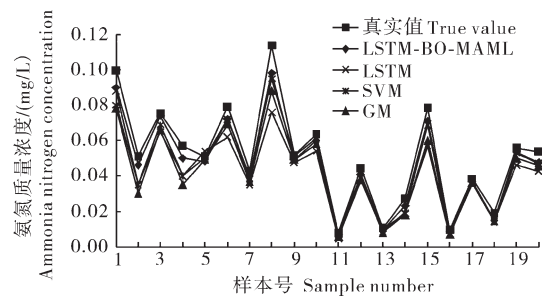


图8 常用少样本学习算法预测结果

Fig.8 Prediction results of commonly used few sample learning algorithms

### 2.4 模型评价

模型评价采用均方根误差(RMSE)、平均绝对误差(MAE)、均方误差(MSE),公式如下:

$$X_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (17)$$

$$X_{MAE} = \frac{\sum_{i=1}^n |\hat{y}_i - y_i|}{n} \quad (18)$$

$$X_{MSE} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (19)$$

式(17)~(19)中: $y_i$ 是真实值; $n$ 是样本数量。

在预测值以及真实值之间,这些指标主要是从误差的比例以及大小来对模型进行分析,可以使预测误差特征得到整体体现。通过这3个指标来对本研究提出的预测模型进行评估,各模型的具体评价结果见表3。

表3中的数据是多次实验结果的平均值,本研究模型在少量数据的训练下,同时利用贝叶斯优化获取超参数以及元学习优化模型参数, RMSE与MAE有所下降,氨氮浓度预测的准确度提高。从表3可以看出,单独使用贝叶斯优化的模型(LSTM-BO)均方

表3 各模型的评价结果

Table 3 Evaluation results of the models

模型 Model	均方根误差 RMSE	平均绝对误差 MAE	均方误差 MSE
GM	0.028 1	0.024 1	0.000 79
SVM	0.027 8	0.023 7	0.000 77
RNN	0.035 3	0.028 9	0.001 20
GRU	0.031 6	0.025 9	0.000 99
LSTM	0.031 1	0.025 5	0.000 96
LSTM-BO	0.029 9	0.024 9	0.000 89
LSTM-MAML	0.027 9	0.024 2	0.000 78
LSTM-BO-MAML	0.027 6	0.023 9	0.000 76

根误差较 LSTM 模型降低了 3.85%，单独使用元学习优化的模型 (LSTM-MAML) 均方根误差较 LSTM 模型降低了 10.2%，本研究所建模型 (LSTM-BOMAML) 均方根误差较 LSTM 模型降低了 11.25%，表明本研究所提出的模型比 LSTM 模型、单独使用贝叶斯和元学习的模型预测效果更好。

与常用小样本学习算法灰色预测模型 (GM) 和支持向量机 (SVM) 进行模型预测结果对比，可以看出，几种模型在少样本数据下预测精度相差不大，灰色预测模型 (GM) 均方根误差较 LSTM 模型降低了 9.64%，支持向量机 (SVM) 均方根误差较 LSTM 模型降低了 10.6%，本文所建模型均方根误差较 LSTM 模型降低了 11.25%。但本研究提出的模型通过较少迭代次数就达到收敛效果，模型误差较低，更适用于长期模型的建立与修正。与常用深度网络 RNN、GRU 模型进行对比评价，LSTM 模型的均方根误差低于 RNN 模型和 GRU 模型，本研究提出的模型误差较 RNN 模型、GRU 模型和 LSTM 模型均有所下降，表明预测效果更好。

### 3 讨论

由于海水养殖中水质环境复杂，对预测准确度要求严格，同时在数据不足的情况下，模型无法做到精确的预测，本研究提出了基于贝叶斯与元学习的氨氮浓度预测模型。相比于其他预测模型，LSTM-BO-MAML 模型既有快速的收敛能力又有较高的预测精度，通过对 LSTM 模型超参数和模型参数进行优化，提高了循环神经网络模型的预测精度和泛化能力，通过较少迭代次数达到收敛效果，模型训练 30 次后损失函数趋于平稳。本研究提出的方法更加适合于小样本数据，然而水质环境的影响因素还有很多，如何进一步获取其他可测因子，提高模型精度和

模型优化的改进方法也是下一步要研究的课题。本研究所提出的建模和优化方法主要针对水产养殖过程，能否拓展到其他领域还需要进一步验证。

### 参考文献 References

- [1] 王静香, 李纪岳, 李树君, 等. 我国水产养殖标准化探索[J]. 中国水产, 2019(4): 31-34. WANG J X, LI J Y, LI S J, et al. Exploration on standardization of aquaculture in China[J]. China fisheries, 2019(4): 31-34 (in Chinese).
- [2] 张成林, 徐皓, 王世明, 等. 基于大型渔业平台的深远海渔业发展现状与思考[J]. 中国农学通报, 2020, 36(25): 152-157. ZHANG C L, XU H, WANG S M, et al. Offshore fisheries based on large-scale fishery platform: development status[J]. Chinese agricultural science bulletin, 2020, 36(25): 152-157 (in Chinese with English abstract).
- [3] 朱日同. 集约化养殖水体氨氮危害及调控措施[J]. 河南水产, 2018(3): 5-7. ZHU R T. The harm and controlling measures of ammonia nitrogen in intensive aquaculture water[J]. Henan fisheries, 2018(3): 5-7 (in Chinese with English abstract).
- [4] 刘桂芳. 水产养殖中氨氮和亚硝酸盐氮的危害及防治[J]. 养殖与饲料, 2018(5): 37-38. LIU G F. Harm and prevention of ammonia nitrogen and nitrite nitrogen in aquaculture[J]. Animals breeding and feed, 2018(5): 37-38 (in Chinese).
- [5] 王魏, 郭戈. 基于随机配置网络的海水养殖氨氮浓度软测量模型[J]. 农业机械学报, 2020, 51(1): 214-220. WANG W, GUO G. Soft measurement model for ammonia nitrogen concentration in marine aquaculture based on stochastic configuration networks[J]. Transactions of the CSAM, 2020, 51(1): 214-220 (in Chinese with English abstract).
- [6] ZHOU J E, QI G H, LIU C Z. A chaotic parallel artificial fish swarm algorithm for water quality monitoring sensor networks 3D coverage optimization[J/OL]. Journal of sensors, 2021, 2021: 5529527 [2022-09-30]. <https://doi.org/10.1155/2021/5529527>.
- [7] DEGHAN-SHOAR M H, ORSI A A, PULLANAGARI R R, et al. A hybrid model to predict nitrogen concentration in heterogeneous grassland using field spectroscopy[J/OL]. Remote sensing of environment, 2023, 285: 113385 [2022-09-30]. <https://doi.org/10.1016/j.rse.2022.113385>.
- [8] 李先鹏, 吴若男, 王魏, 等. 基于小波降噪和 LSTM 的海参养殖氨氮预测[J]. 控制工程, 2022, 29(4): 587-592, 626. LI X P, WU R N, WANG W, et al. Prediction of ammonia nitrogen concentration in sea cucumber aquaculture water based on wavelet denoising and LSTM[J]. Control engineering of China, 2022, 29(4): 587-592, 626 (in Chinese with English abstract).
- [9] 冯良记, 唐军. 改进遗传算法在高非线性水质模型参数估值中的应用研究[J]. 水利与建筑工程学报, 2009, 7(2): 11-13. FENG L J, TANG J. Research on application of modified genetic algorithm in parameter estimation of highly non-linear water quality model[J]. Journal of water resources and architectural engineering, 2009, 7(2): 11-13 (in Chinese with English abstract).

- [10] 唐亦舜,徐庆,刘振鸿,等.基于优化非线性自回归神经网络模型的水质预测[J].东华大学学报(自然科学版),2022,48(3):93-100.TANG Y S,XU Q,LIU Z H, et al.Prediction of water quality based on the optimized nonlinear auto-regressive neural network model[J].Journal of Donghua University (natural science edition),2022,48(3):93-100(in Chinese with English abstract).
- [11] 王魏,邓长辉,关晓燕,等.水产养殖溶氧浓度软测量模型研究进展[J].渔业现代化,2015,42(6):26-30.WANG W,DENG C H,GUAN X Y, et al.Research progress on the soft-sensing model for dissolved oxygen in aquaculture[J].Fishery modernization,2015,42(6):26-30 (in Chinese with English abstract).

## Optimization of ammonia concentration prediction model based on Bayesian and Meta-learning

LIU Yiwei<sup>1</sup>, WANG Wei<sup>1</sup>, ZHANG Shuyu<sup>1</sup>, SUN Junyang<sup>1</sup>, LI Shuangshuang<sup>2</sup>

1.School of Information Engineering, Dalian Ocean University/Key Laboratory of Environment Controlled Aquaculture, Ministry of Education, Dalian 116023, China;  
2.Dalian Xinyulong Marine Biological Seed Technology Co. Ltd., Dalian 116000, China

**Abstract** To address the issues of low prediction accuracy and slow convergence rate of ammonia nitrogen concentration model under small sample data, a method of establishing ammonia nitrogen concentration prediction model by using long short-term memory (LSTM) and optimizing the model by using Bayesian optimization algorithm and Meta-learning mechanism was proposed. The Bayesian optimization algorithm was used to optimize the superparameters of the prediction model, and the initial values of the model parameters were given. Then the Meta-LSTM algorithm was used to learn the model gradient and allow the parameter sharing and updating among the optimizers, and finally the optimization of the prediction model of ammonia nitrogen concentration was realized. Compared with LSTM, GRU and RNN models, the result shows that the root-mean-square error, mean absolute error and mean square error of the proposed model are 0.027 6, 0.023 9 and 0.000 76, respectively, which are better than other prediction models. It is further indicated that the prediction model of ammonia nitrogen concentration based on Bayesian and Meta-learning is effective for modeling small sample data, and can achieve convergence in rapid training. The accuracy of the model meets the prediction requirements of ammonia nitrogen concentration for aquaculture.

**Keywords** Bayesian optimization; ammonia nitrogen concentration prediction; Meta-LSTM; Meta-learning

(责任编辑:边书京)