

基于LSTM网络的海水温度剖面预报研究

范培勤¹, 过武宏^{1*}, 唐帅¹, 张驰¹, 曲泓玥²

(1. 海军潜艇学院, 山东 青岛 266199; 2. 青岛海洋科学与技术试点国家实验室, 山东 青岛 266000)

摘要: 基于海水温度历史观测和海洋模式数值预报数据, 利用长短期记忆(LSTM)神经网络开展了海水温度剖面短时预报方法研究。以南海17°46.91'N, 112°03.24'E处海水温度剖面预报为研究对象, 利用模式预报数据和观测数据, 构建了观测、预报、观测和预报混合3个样本数据集。基于LSTM神经网络模型, 建立了由编码器-解码器组成的多对多海水温度剖面序列预报模型, 并开展了模型的训练和性能验证分析工作。结果表明: 该模型具有较高的预报精度, 对小样本问题的处理具有良好的稳定性; 采用深度分层预报, 可有效提高模型预报精度, 改善模型泛化能力; 与只使用观测数据集的预报相比, 混合样本集预报误差下降明显, 为基于小样本观测数据的海洋环境要素预报研究提供了一种思路。

关键词: 深度学习; 长短期记忆神经网络; 海水温度; 短时预报

中图分类号: P731.31 **文献标识码:** A **文章编号:** 1003-0239(2024)03-0033-11

0 引言

实时、动态、精确的海洋环境信息对优化水声传感器的使用、提高水下攻防作战效能具有重要意义。近几十年来, 海洋数值模式和海洋观测均取得了长足的进步, 数值模式预报逐渐成为海洋环境信息预报和保障的主要手段, 预报结果的种类、范围、精细化程度都有了大幅提升^[1-3], 但受数值模式近似计算和观测资料测量误差等因素的影响, 海洋数值模式的预报结果难以水下作战平台提供准确的实时或临近预报信息支持^[4]。数值预报和海洋观测的快速发展, 积累了大量预报和观测数据, 为利用神经网络技术进行海洋环境短时预报研究提供了良好的数据基础。如何利用深度学习对海洋环境数据进行充分的价值挖掘, 探索数据间内在关联, 实现海洋环境保障信息的实时化和精确化预报, 逐渐成为研究热点^[5]。

海水温度剖面数据是典型的时间序列数据, 循

环神经网络(Recurrent Neural Networks, RNN)是一类具有短期记忆能力的神经网络, 适用于处理时间序列预测问题, 被广泛应用于语音识别、情感分类和自然语言处理等任务中。针对RNN在处理较长时间序列时存在的梯度消失和爆炸问题, 长短期记忆(Long Short-Term Memory, LSTM)神经网络于1997年被提出, LSTM单元可以学会识别重要的输入, 将其存储在长期状态中, 并在需要的时候提取出来, 防止较早信息在处理过程中逐渐消失, 在长时间、多变量、多输入预报问题的处理中取得了较好的结果^[6-7]。LSTM深度神经网络可用于开展气温、地温、降水等气象要素预测的研究^[8-12]以及海表温度、沿海风速、海冰、西太平洋暖池变化等海洋环境要素的预报工作^[13-17]。本文基于长时间海水温度剖面预报和观测资料, 利用LSTM神经网络, 构建了基于编码器-解码器结构的海水温度剖面预报模型, 开展了基于观测数据、数值预报数据、观测和预报混合数据的多时间步长海水温度剖面预报的研

收稿日期: 2022-07-28。

基金项目: 基础计划加强重点基础研究项目(2020-JCJQ-ZD-144-00, 2020-JCJQ-ZD-144-03)。

作者简介: 范培勤(1981-), 男, 副研究员, 博士, 主要从事水声环境效应研究。

*通信作者: 过武宏(1980-), 男, 副教授, 博士, 主要从事水声环境效应研究。

究工作,并对计算结果进行了对比分析。

1 数据处理

1.1 数据来源

分别选取南海海区17°46.91'N,112°03.24'E处的海洋模式预报、海上观测的海水温度剖面数据开展相关的工作,其中,预报数据时间跨度为2020年3月23日—2022年2月21日,观测数据为2020年11月25日—2021年8月4日(见表1)。

表1 样本数据集信息统计

数据类型	深度范围/m	垂直层数	时间间隔	预报时效	天数/d	数据量
预报	0~2 000	42	6 h	168 h	700	2 800
观测	110~1 800	23	1 min	实时	248	357 120

1.2 数据预处理

1.2.1 数据标准化

①观测数据集

与预报数据相比,观测数据受周边环境因素影响较大,例如观测阵受海浪、海流等因素影响会存在摆动,造成传感器观测深度偏离预设深度,影响观测数据序列的一致性,为提高数据的可用性,须对观测数据进行标准化处理。经统计,所有观测数据序列水深的交集为190~1 500 m,由于数据缺失严重,难以通过插值等方式在深度方向补齐。为便于后续分析比较,只对190~1 500 m的观测数据通过线性插值进行标准化处理。为比较LSTM模型对不同时间步长温度剖面的预报能力,分别按时间步长1 min、1 h、6 h整理出3组观测数据集。标准化处理后观测数据集信息见表2。

②预报数据集

模式预报结果通常进行过质量控制,数据质量较好,故只在时间轴上对缺失数据进行补齐即可。取每天逐6 h预报结果中的前4组数据作为当天4个时刻的预报结果,空值和缺失值取前一天预报结果中相应时段的数值补齐,若连续缺失天数超出上一条数据的预报时效,则使用上一条数据与下一条

表2 观测数据集信息统计

Tab.2 Observation data set information

观测数据集	深度范围/m	垂直层数	时间步长	数据量
1	190~1 500	12	1 min	357 120
2	190~1 500	12	1 h	5 952
3	190~1 500	12	6 h	992

数据间的线性插值补齐。为便于分析在预报和观测数据集合并后,模型预报误差的变化情况,将预报数据水深统一到1 500 m。标准化处理后预报数据集信息见表3。

表3 预报数据集信息统计

Tab.3 Forecasting data set information

预报数据集	深度范围/m	垂直层数	时间步长/h	数据量
1	0~1 500	42	6	2 800
2	190~1 500	12	6	2 800

1.2.2 数据的归一化处理

为提高模型的收敛速度和稳定性,使用线性函数对样本数据进行归一化处理,使数值介于[0,1]。计算公式为:

$$\bar{X} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

式中: X 为原始数据; \bar{X} 为归一化处理后的数据; X_{\max} 和 X_{\min} 分别为原始数据中的最大值和最小值。

1.2.3 输入/输出序列构建

与卷积等神经网络模型相比,循环神经网络不受输入/输出数据形状的限制,便于实现一对多、多对多等输入/输出序列问题的预报工作。本文利用前4个时刻的海水温度剖面数据预报未来4个时刻的剖面数据,是一种典型的多对多预报问题。输入/输出序列的构建方法如下:假设有 N 个样本,使用前 m 个时刻的数据 $X_i = \{x_1, x_2, \dots, x_m\}$ 预报未来 n 个时刻的结果 $Y_i = \{x_{m+1}, x_{m+2}, \dots, x_{m+n}\}$,即建立其中 m 个时间点样本与未来 n 个时间点样本之间的映射关系,将它们分别作为网络的输入参数和期望输出结果进行训练。在训练结束后,用得到的模型进行预报。对于每一个训练样本,本文 m 、 n 均取4。输入和输出序列的结构如下:

$$X = \{X_1, X_2, \dots, X_{N-m-n}\} \quad (2)$$

$$Y = \{Y_1, Y_2, \dots, Y_{N-m-n}\} \quad (3)$$

2 模型构建

时间序列预报分析是利用过去一段时间内某事件时间的特征来预报未来一段时间内该事件的特征。与回归分析模型预报不同,时间序列预报分析处理的是输入序列和输出序列之间的映射关系,不再是单个数值之间的映射,时间序列依赖于事件发生的先后顺序,具有时间相关性。

2.1 模型搭建

Seq2Seq (Sequence to Sequence) 模型可以很好地解决输入/输出序列长度不一致的问题。模型由一个编码器与一个解码器组成,编码器将输入序列映射为特定长度的向量,学习输入序列中步骤之间的关系,并开发这些关系的内部表示形式,输出一个固定大小的向量,代表输入序列的内部表示;解码器使用来自编码器的状态向量作为初始状态,并将输入序列的内部表示形式转换为正确的输出序列。具体构建过程为^[7]:

①编码器构建:编码器子模型可以使用一层或多层LSTM层实现,本文选用一层LSTM网络,神经元数量为64个。

②时间步长重复:由于输入序列具有多个时间步长,需对输入序列内部每个时间步长重复一次,重复后的向量序列将被提交给LSTM解码。Repeatvector()作为适配器,将编码器和解码器连接在一起。

③解码器:一个或多个LSTM层都可以用于实现解码器模型,该模型从编码器模型的固定大小输出中读取,将学习到的输入序列内部表示形式转换为正确的输出序列。本文选用一层LSTM网络,神经元数量为64个。

④输出层:使用一个全连接层输出序列中的每个时间步。输出层预报输出序列中的一个步骤,而不是一次几个时间步长。这意味着解码器提供的每个时间步将使用相同的全连接层和输出层来处理。为了实现这一点,把解码器和输出层包装在Timedistributed()包装器中,用于解码器的每个时间

步输出。

整个模型的计算流程见图1。

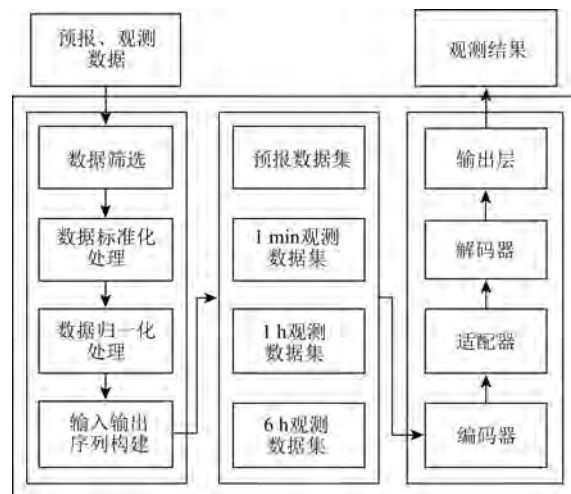


图1 LSTM模型计算流程图

Fig.1 Flow chart of the LSTM model

2.2 模型性能分析

本文采用均方根误差(Root Mean Square Error, RMSE)、平均绝对误差(Mean Absolute Error, MAE)对模型预报性能进行评估。计算公式为:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N-m-n} (\hat{Y}_i - Y_i)^2}{N-m-n}} \quad (4)$$

$$MAE = \frac{\sum_{i=1}^{N-m-n} |\hat{Y}_i - Y_i|}{N-m-n} \quad (5)$$

式中: N 为样本总数目; m 为输入时间步长; n 为输出时间步长; Y_i 为验证集第 i 个输入变量 X_i 对应的真值; \hat{Y}_i 为验证集第 i 个输入变量对应的输出值。RMSE和MAE值越低,表示模型预报精度越高。

3 结果分析

3.1 测试方法

为了验证LSTM模型的预报能力,分别利用观测、预报、观测和预报混合数据集对模型进行训练和测试。在典型的热带和温带深海环境下,温度的垂直剖面通常可以分为混合层、温跃层和深海等温

层。混合层由于受海面的冷热交换和风浪搅拌作用,在海表以下一定深度内的海水温度基本不变,该层水温的日变化较大;混合层下为温跃层,该层随水深的增加,海水温度急剧下降,除靠近混合层的深度外,该层水温受季节变化影响不大;随着水深的继续增加,到一定深度后海水上下层热量交换基本达到平衡,海水温度基本保持不变直到海底,形成深海等温层。该研究海区深海等温层大约从水下 1 000 m 开始直到海底^[18]。由于数据集的取值范围对模型预报精度有较大的影响,本文按照温度剖面垂直结构特点,将温度剖面划分为若干子剖面,以此来控制样本数据的取值范围,然后按照数据未分层、分层两种情况分别开展训练和验证工作。垂直深度上分层的原则和步骤为:

第一步:根据深海温度剖面在深度方向上垂直结构特点(混合层 ≤ 50 m、温跃层 $\leq 1\,000$ m和深海等温层 $\geq 1\,000$ m)划分为3个子集。

第二步:利用建立的模型对3个子集分别进行测试,按照预报结果绝对误差 $MAE \leq 0.5\text{ }^{\circ}\text{C}$ 的原则^[9],将每个子集再划分为若干子集。具体划分方法为:首先将整个子集作为样本提供给模型进行训练验证,若验证结果绝对误差大于 $0.5\text{ }^{\circ}\text{C}$,则将子集最后一层去掉重新生成一个子集提交模型进行训练和验证,如此循环,直到新划分子集验证结果绝对误差小于 $0.5\text{ }^{\circ}\text{C}$ 后,完成一个新子集的划分。然后将剩余样本按照上述方法继续进行划分,最终得到新的样本子集。

按上述原则和步骤,最终得到的样本子集划分情况见表4。

表 4 数据子集划分情况统计表

Tab.4 Data subset division

预报数据集		观测数据集	
水深/m	层数	水深/m	层数
1~20	6	—	—
25~50	6	—	—
55~80	6	—	—
85~120	6	—	—
130~180	6	—	—
190~250	4	190~250	4
275~400	4	275~400	4
500~800	2	500~800	2
1 000~1 500	2	1 000~1 500	2

3.2 模式预报数据集测试结果

预报数据样本集及其子集各有 2 800 条数据,将其前 2 404 条做训练集,后 396 条做验证集,分别采用滑窗方式得到时间序列训练样本 2 400 条、验证集 392 条,开始模型的训练验证工作。表 5 为采用未分层、分层两种预报样本数据集训练得到的模型验证结果误差,其中 MAE_{\max} 和 $RMSE_{\max}$ 分别为该水层 4 个时间步长预报结果中的最大绝对误差值和最大均方根误差值。

表 5 预报数据验证集误差统计

Tab.5 Error satistics of forecasting data verification set

水深/m	层数	$T_{\max}/^{\circ}\text{C}$	$T_{\min}/^{\circ}\text{C}$	$\Delta T/^{\circ}\text{C}$	未分层预报结果/ $^{\circ}\text{C}$		分层预报结果/ $^{\circ}\text{C}$	
					MAE_{\max}	$RMSE_{\max}$	MAE_{\max}	$RMSE_{\max}$
1~20	6	31.10	23.27	7.83	0.449	0.243	0.081	0.015
25~50	6	30.42	22.92	7.50	0.525	0.313	0.076	0.009
55~80	6	27.26	19.67	7.59	0.694	0.582	0.192	0.055
85~120	6	24.76	17.06	7.70	0.415	0.244	0.397	0.222
130~180	6	20.90	14.42	6.48	0.687	0.648	0.184	0.077
190~250	4	16.30	12.22	4.08	0.823	0.751	0.099	0.018
275~400	4	13.19	9.42	3.77	0.447	0.277	0.092	0.012
500~800	2	9.20	5.22	3.98	0.229	0.068	0.027	0.001
1 000~1 500	2	4.79	2.99	1.80	0.123	0.021	0.009	<0.001

从表5中可以看出,未分层、分层两种情况的预报结果最大绝对误差分别为 $0.823\text{ }^{\circ}\text{C}$ 、 $0.397\text{ }^{\circ}\text{C}$,分层情况下误差降低50%以上;最大均方根误差分别为 $0.751\text{ }^{\circ}\text{C}$ 、 $0.222\text{ }^{\circ}\text{C}$,分层情况下误差降低约70%;分层后的最大绝对误差、最大均方根误差均小于 $0.5\text{ }^{\circ}\text{C}$,说明分层预报结果明显优于未分层预报。

图2为未分层、分层两种情况下预报结果的绝对误差、均方根误差时空分布。从图中可以看出,在两种情况下,200 m以浅深度水温预报误差较大且波动剧烈,200 m以深预报误差较小且变化平缓,这主要是由于海表附近受太阳照射、风浪搅拌等因素的影响,水温变化较大引起的。分层预报误差比未分层预报误差整体大幅下降,且深度上变化更加平稳。

3.3 观测数据集测试结果

本节用表1中时间间隔分别为1 min、1 h、6 h的3个观测数据样本集,针对未分层和分层两种情况对模型进行训练和测试,研究模型随训练样本数量

减小的性能变化情况。

3.3.1 1 min观测样本集测试结果

1 min观测数据样本集共有357 120条数据,将前320 004条做训练集,后37 116条做验证集,分别采用滑窗方式得到时间序列训练样本32 000条、验证集37 112条,利用模型开始训练验证工作。表6为在未分层、分层情况下,由1 min观测样本数据集训练得到的模型验证结果误差。

从表6中可以看出,该模型具有较好的预报准确性,在分层和未分层两种情况下,两种方法的预报结果误差相差不大,整个水深上最大绝对误差 $\leq 0.054\text{ }^{\circ}\text{C}$ 、均方根误差 $<0.006\text{ }^{\circ}\text{C}$,均远小于 $0.5\text{ }^{\circ}\text{C}$,两种情况下模型均具有较高的预报精度。

图3为两种情况下预报结果的绝对误差、均方根误差时空分布。从图中可以看出,模型具有较高的预报精度,预报误差随水深增加变化平稳。

3.3.2 1 h实测样本集测试结果

1 h观测数据样本集共有5 952条数据,将前

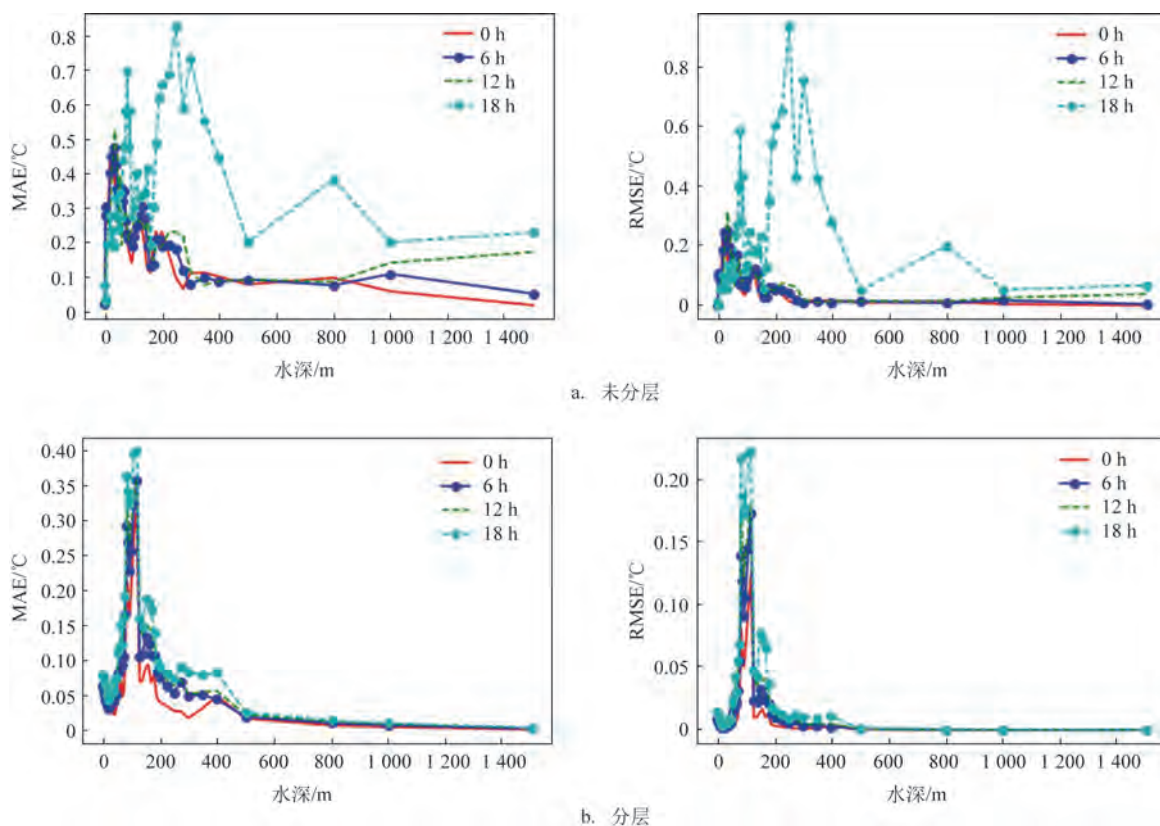


图2 预报数据验证集误差时空分布

Fig.2 Spatial-temporal distribution error of forecast data verification set

表6 1 min观测数据验证集误差统计

Tab.6 Error statistics of 1 min observation data verification set

水深/m	层数	$T_{\max}/^{\circ}\text{C}$	$T_{\min}/^{\circ}\text{C}$	$\Delta T/^{\circ}\text{C}$	未分层预报结果/ $^{\circ}\text{C}$		分层预报结果/ $^{\circ}\text{C}$	
					MAE _{max}	RMSE _{max}	MAE _{max}	RMSE _{max}
190~250	4	17.12	11.25	0.55	0.054	0.006	0.054	0.006
275~400	4	13.68	8.89	0.27	0.030	0.002	0.030	0.002
500~800	2	9.36	5.09	0.16	0.014	<0.001	0.014	<0.001
1 000~1 500	2	4.80	3.00	0.10	0.006	<0.001	0.006	<0.001

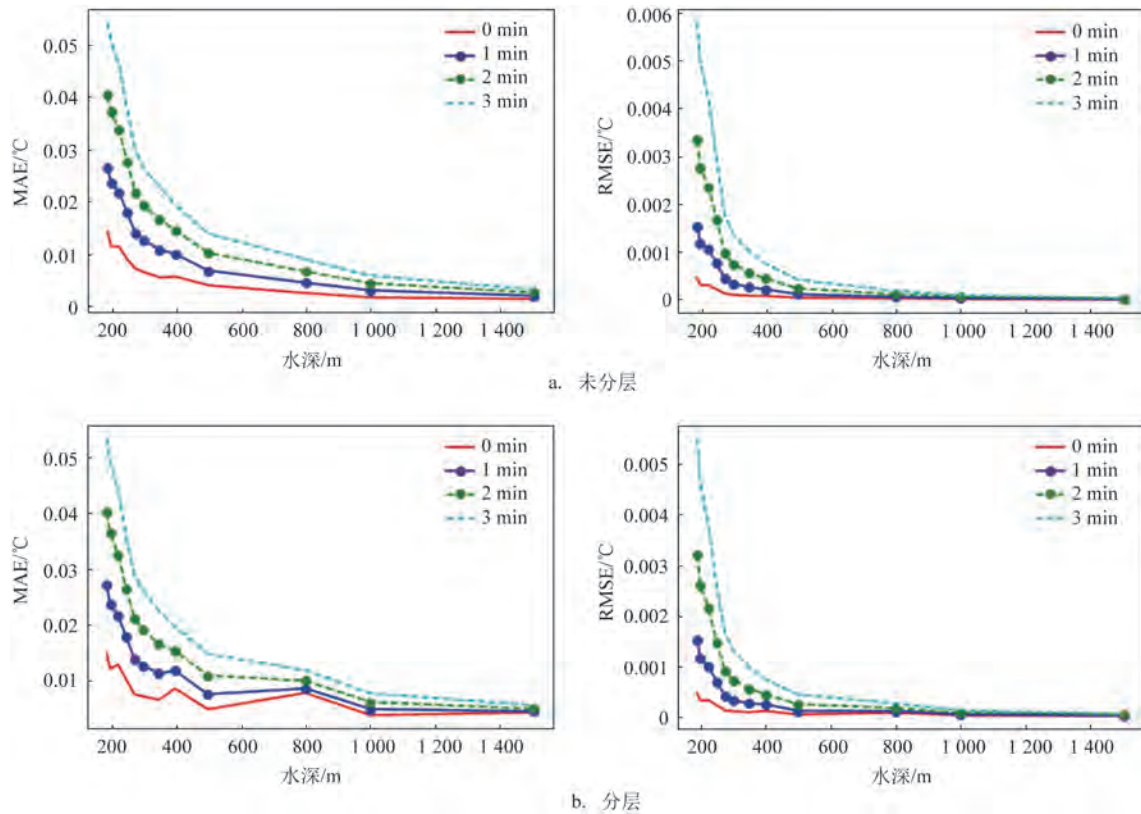


图3 1 min观测样本验证集误差时空分布

Fig.3 Spatial-temporal distribution error of 1 min observation data verification set

5 404条做训练集,后548条做验证集,分别采用滑动窗口方式得到时间序列训练样本5 400条、验证集544条,利用模型开始训练验证工作。表7为在未分层、分层情况下,由1 h观测样本数据集训练得到的模型验证结果误差。

从表中可以看出,随着样本数量的减少,预报误差显著增加,未分层、分层两种情况的预报结果最大绝对误差分别为0.353 $^{\circ}\text{C}$ 、0.250 $^{\circ}\text{C}$,分层情况下误差降低约30%;最大均方根误差分别为

0.204 $^{\circ}\text{C}$ 、0.097 $^{\circ}\text{C}$,分层情况下误差降低50%以上;分层后的最大绝对误差、最大均方根误差均小于0.5 $^{\circ}\text{C}$,分层预报结果明显优于未分层预报结果。

图4为未分层、分层两种情况下预报结果的绝对误差、均方根误差时空分布。从图中可以看出,分层预报结果误差比未分层预报下降明显,且变化更加平稳。

3.3.3 6 h实测样本集测试结果

6 h观测数据样本集共有992条数据,将前904

表7 1 h观测数据验证集误差统计

Tab.7 Error statistics of 1 h observation data verification set

水深/m	层数	$T_{\max}/^{\circ}\text{C}$	$T_{\min}/^{\circ}\text{C}$	$\Delta T/^{\circ}\text{C}$	未分层预报结果/ $^{\circ}\text{C}$		分层预报结果/ $^{\circ}\text{C}$	
					MAE _{max}	RMSE _{max}	MAE _{max}	RMSE _{max}
190~250	4	17.12	11.27	5.85	0.353	0.204	0.250	0.097
275~400	4	13.45	8.91	4.54	0.220	0.072	0.179	0.049
500~800	2	9.33	5.21	4.12	0.093	0.014	0.079	0.012
1 000~1 500	2	4.80	3.03	1.77	0.059	0.006	0.078	0.009

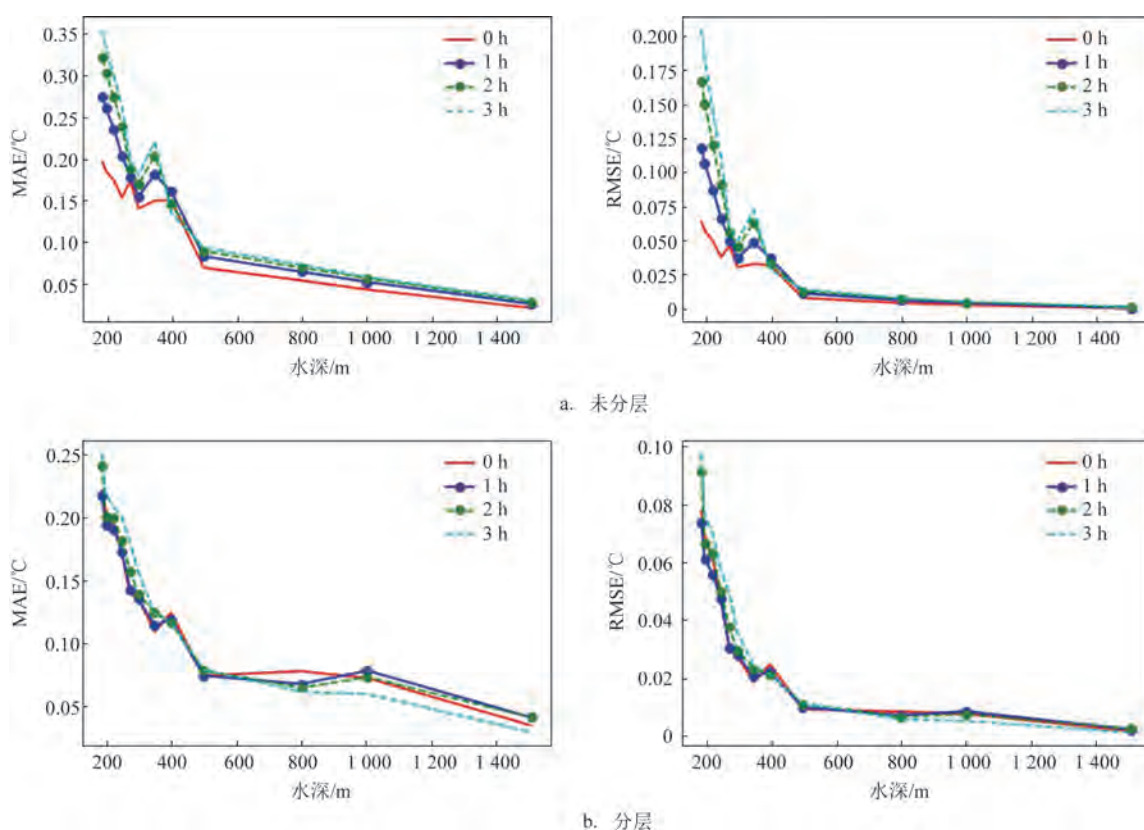


图4 1 h观测数据验证集误差时空分布

Fig.4 Spatial-temporal distribution error of 1 h observation data verification set

条做训练集,后88条做验证集,分别采用滑窗方式得到时间序列训练样本900条、验证集84条,利用模型开始训练验证工作。表8为未分层、分层情况下,由6 h观测样本数据集训练得到的模型验证结果误差。

从表中可知,在未分层、分层两种情况下,整个水深的预报结果的最大绝对误差分别为0.324 $^{\circ}\text{C}$ 、0.271 $^{\circ}\text{C}$,分层情况下误差降低约16%;最大均方根

误差分别为0.149 $^{\circ}\text{C}$ 、0.098 $^{\circ}\text{C}$,分层情况下误差降低约35%;分层后的最大绝对误差、最大均方根误差均小于0.5 $^{\circ}\text{C}$,分层预报结果优于未分层预报结果。

图5为未分层、分层两种情况下预报结果的绝对误差、均方根误差时空分布。从图中可以看出,在两种情况下,预报误差在1 000 m以浅水深上变化较大,1 000 m以深变化平稳;分层预报误差比未分层预报误差下降明显,且在深度上变化更加平稳。

表8 6 h观测数据验证集误差统计

Tab.8 Error statistics of 6 h observation data verification set

水深/m	层数	$T_{\max}/^{\circ}\text{C}$	$T_{\min}/^{\circ}\text{C}$	$\Delta T/^{\circ}\text{C}$	未分层预报结果/ $^{\circ}\text{C}$		分层预报结果/ $^{\circ}\text{C}$	
					MAE _{max}	RMSE _{max}	MAE _{max}	RMSE _{max}
190~250	4	16.99	11.50	5.49	0.324	0.149	0.271	0.098
275~400	4	13.43	8.95	4.48	0.278	0.101	0.162	0.043
500~800	2	9.22	5.26	3.96	0.222	0.021	0.117	0.019
1 000~1 500	2	4.75	3.07	1.68	0.074	0.008	0.049	0.004

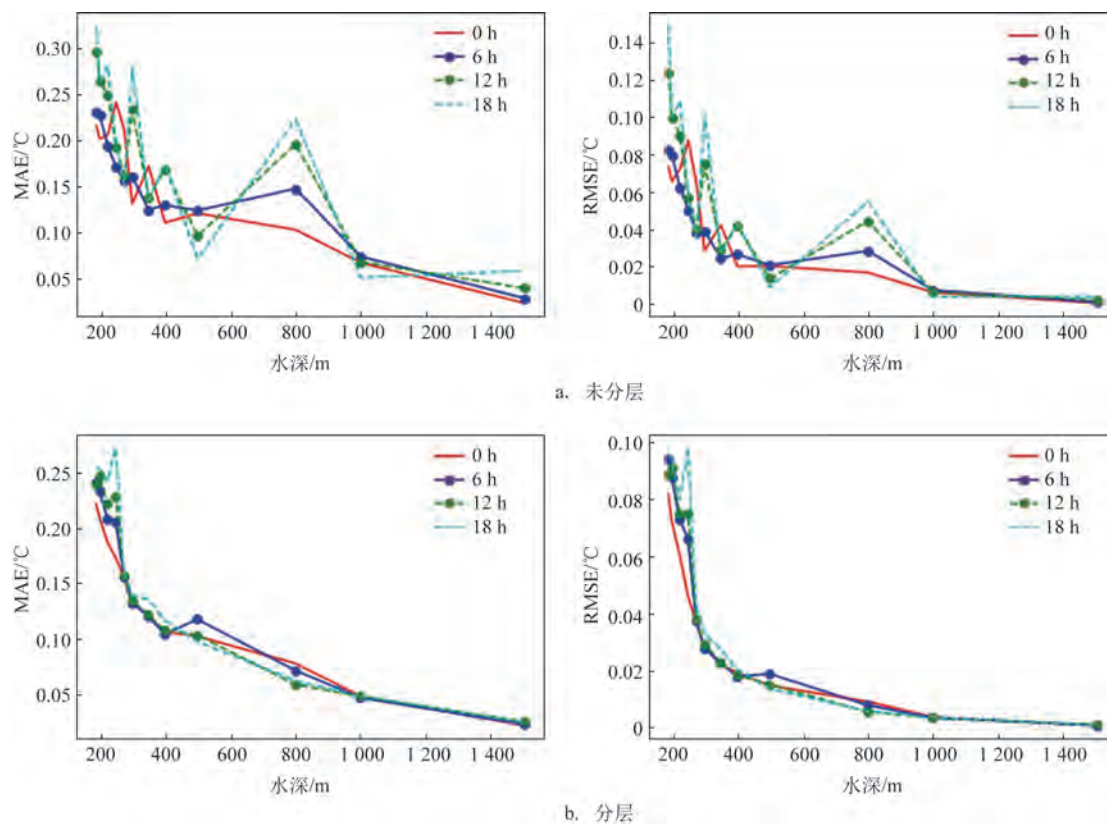


图5 6 h观测数据验证集误差时空分布

Fig.5 Spatial-temporal distribution error of 6 h observation data verification set

3.4 模式预报和观测混合数据集测试结果

由于长时间、大范围的海水温度剖面观测数据获取难度大,基于观测数据的海水温度剖面预报是一个典型的小样本学习问题。本文针对观测数据样本集规模小、获取难度大的问题,将模式预报数据集加入到6 h观测数据集中,探索利用模式预报数据扩充样本数据量、提高小样本情况下模型预报准确性这一方法的可行性。

具体做法为:6 h观测数据样本集共有992条数据,将前804条做训练集,后198条做验证集,分别采用滑窗方式得到时间序列训练样本800条、验证集194条;预报数据集共有2 800条数据,采用滑窗方式得到时间序列训练样本2 796条。将800条实测数据、2 796条预报数据样本合并作为模型的训练数据,将194条验证数据作为测试集,利用模型开始训练验证工作。表9为采用未分层、分层情况下,由6 h观测、预报混合数据验证集训练得到的模型验证

结果误差。

从表9中可以看出,在未分层、分层两种情况下,模型预报结果的最大绝对误差分别为0.245℃、0.236℃,两者相差不大;最大均方根误差分别为0.097℃、0.072℃,分层情况下误差降低25%以上;分层后的最大绝对误差、最大均方根误差均小于0.5℃,分层预报结果优于未分层预报结果。

由表8、表9可以看出,与只使用观测数据相比,

使用混和样本数据训练得到的模型预报误差降低明显,在未分层、分层两种情况下,最大绝对误差分别下降约24%、13%,最大均方根误差分别下降约35%、27%。图6为两种情况下预报结果的绝对误差、均方根误差时空分布,从中可以看出,500 m以浅的预报误差波动较大,这主要由于海表附近温度变化较大引起,500 m以深由于温度变化平缓,预报误差变化不大。由此可见,若样本数据存在较大波动,会对预报结果产生明显影响。

表9 6 h观测、预报混合数据验证集误差统计

Tab.9 Error Statistics of 6 h mixed observation-forecast data verification set

水深/m	层数	$T_{\max}/^{\circ}\text{C}$	$T_{\min}/^{\circ}\text{C}$	$\Delta T/^{\circ}\text{C}$	未分层预报结果/ $^{\circ}\text{C}$		分层预报结果/ $^{\circ}\text{C}$	
					MAE _{max}	RMSE _{max}	MAE _{max}	RMSE _{max}
190~250	4	16.99	11.93	5.06	0.245	0.097	0.236	0.072
275~400	4	13.43	8.95	4.48	0.230	0.071	0.173	0.046
500~800	2	9.22	5.22	4.00	0.125	0.021	0.119	0.022
1 000~1 500	2	4.80	3.00	1.80	0.148	0.026	0.052	0.004

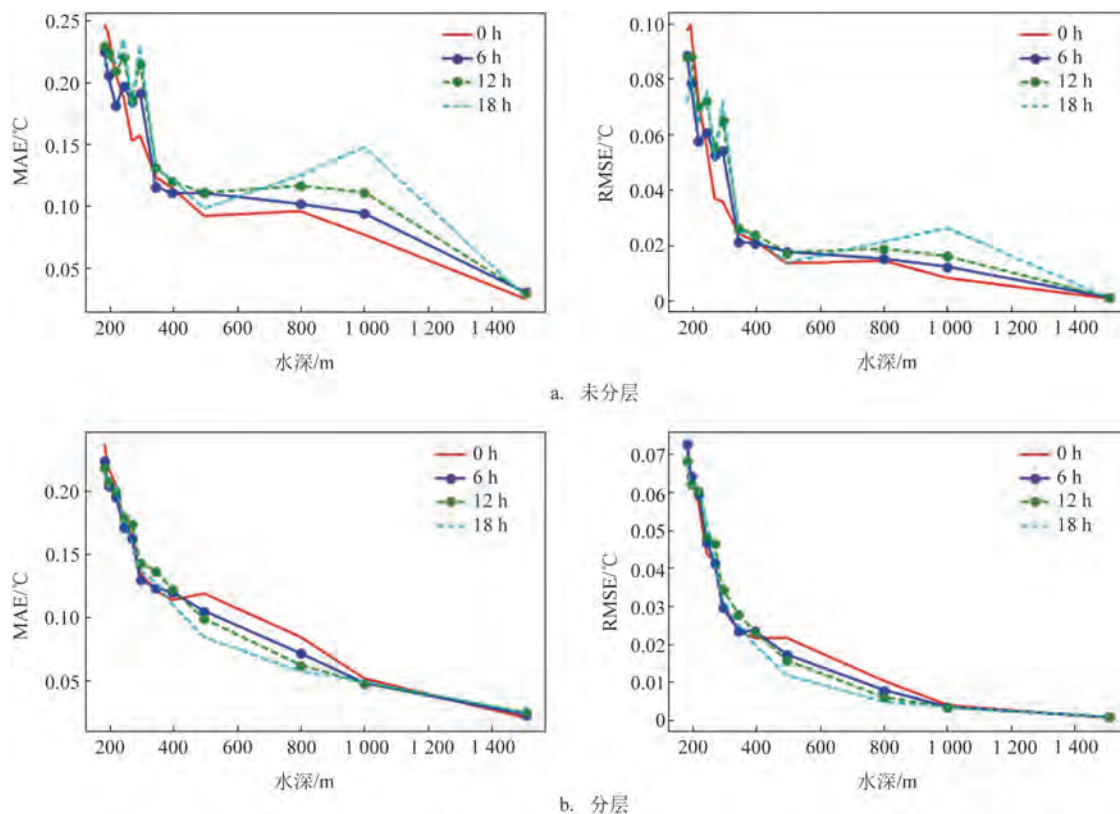


图6 6 h观测、预报混合数据验证集误差时空分布

Fig.6 Spatial-temporal distribution error of 6 h observation data verification set

3.5 小结

通过以上分析可以得出以下结论:

① LSTM模型可以很好地处理海水温度剖面预报问题,对小样本问题也有较好的预报能力。

②样本量的大小对模型的预报精度有显著的影响,样本量减小,训练得到的模型预报精度明显下降。

③预报结果误差与海水温度在深度上的变化相关,温度变化越大,误差越大,采用分层预报可大幅降低预报误差。

4 结论

本文利用海水温度剖面历史观测和模式预报数据,结合 LSTM深度神经网络模型,开展了基于深度学习的海水温度短期预报方法研究。结论如下:

①LSTM编码器-解码器神经网络模型在海水温度短期预报方面具有良好的性能,对小样本问题具有较好的泛化能力。

②采用观测和预报混合数据集,可充分发挥预报数据数量大且完备的优势,并大幅提升模型对观测数据集的预报精度,为基于小样本观测数据集的海洋环境要素预报提供思路。

③针对预报要素取值范围空间大、变化剧烈的问题,采用分层预报方法可有效提高模型的预报精度。

本文的研究工作只针对一个位置点开展,虽然模型表现出良好的预报精度,但其在三维空间上的局限性也非常明显,后面将开展基于深度学习的典型海区的海洋环境要素预测方法研究,进一步提高海洋环境信息保障的有效性和针对性。

参考文献:

- [1] 王辉,万莉颖,秦英豪,等.中国全球业务化海洋学预报系统的发展和应用[J].地球科学进展,2016,31(10):1090-1104.
WANG H, WAN L Y, QIN Y H, et al. Development and application of the Chinese global operational oceanography forecasting system [J]. *Advances in Earth Science*, 2016, 31(10): 1090-1104.
- [2] 刘娜,王辉,凌铁军,等.全球业务化海洋预报进展与展望[J].地球科学进展,2018,33(2):131-140.
LIU N, WANG H, LING T J, et al. Review and prospect of global operational ocean forecasting[J]. *Advances in Earth Science*, 2018,

33(2): 131-140.

- [3] 张志远,孙立尹,何锡玉,等.海洋环境数值预报业务综合运维平台设计与实现[J].海洋预报,2019,36(3):63-70.
ZHANG Z Y, SUN L Y, HE X Y, et al. Design and implementation of marine environment numerical forecast integrated operation and maintenance platform[J]. *Marine Forecasts*, 2019, 36(3): 63-70.
- [4] 马继瑞,韩桂军,李威,等.海洋三维温盐流数值模拟研究的有关进展和问题[J].海洋学报,2014,36(1):1-6.
MA J R, HAN G J, LI W, et al. The relevant progress and issues in numerical simulations of the oceanic 3-D temperature, salinity and current[J]. *Acta Oceanologica Sinica*, 2014, 36(1): 1-6.
- [5] 宋振亚,刘卫国,刘鑫,等.海量数据驱动下的高分辨率海洋数值模式发展与展望[J].海洋科学进展,2019,37(2):161-170.
SONG Z Y, LIU W G, LIU X, et al. Research progress and perspective of the key technologies for ocean numerical model driven by the mass data[J]. *Advances in Marine Science*, 2019, 37(2): 161-170.
- [6] 刘文如.零基础入门Python深度学习[M].北京:机械工业出版社,2020.
LIU W R. Zero basic introduction to Python deep learning[M]. Beijing: China Machine Press, 2020.
- [7] GÉRON A. 机器学习实战-基于Scikit-Learn, Keras 和 Tensorflow [M]. 宋能辉,李炯,译.北京:机械工业出版社,2020.
GÉRON A. Hands-on machine learning with Scikit-Learn, Keras & TensorFlow[M]. SONG N H, LI X, trans. Beijing: China Machine Press, 2020.
- [8] 倪铮,梁萍.基于LSTM深度神经网络的精细化气温预报初探[J].计算机应用与软件,2018,35(11):233-236.
NI Z, LIANG P. Fine temperature forecast based on LSTM deep neural network[J]. *Computer Applications and Software*, 2018, 35(11): 233-236.
- [9] 智协飞,王田,季焱.基于深度学习的中国地面气温的多模式集成预报研究[J].大气科学学报,2020,43(3):435-446.
ZHI X F, WANG T, JI Y. Multimodel ensemble forecasts of surface air temperature over China based on deep learning approach[J]. *Transactions of Atmospheric Sciences*, 2020, 43(3): 435-446.
- [10] 马景奕,刘维成,闫文君.基于深度学习的气象要素预测方法[J].热带气象学报,2021,37(2):186-193.
MA J Y, LIU W C, YAN W J. Meteorological elements forecasting method based on deep learning[J]. *Journal of Tropical Meteorology*, 2021, 37(2): 186-193.
- [11] 唐旺,马尚昌,李程.基于滑动窗口的LSTM地温预测方法[J].成都理工大学学报(自然科学版),2021,48(3):377-384.
TANG W, MA S C, LI C. LSTM ground temperature prediction method based on sliding window[J]. *Journal of Chengdu University of Technology (Science & Technology Edition)*, 2021, 48(3): 377-384.
- [12] 沈皓俊,罗勇,赵宗慈,等.基于LSTM网络的中国夏季降水预测研究[J].气候变化研究进展,2020,16(3):263-275.

- SHEN H J, LUO Y, ZHAO Z C, et al. Prediction of summer precipitation in China based on LSTM network[J]. *Climate Change Research*, 2020, 16(3): 263-275.
- [13] 王国松, 王喜冬, 侯敏, 等. 基于观测和再分析数据的LSTM深度神经网络沿海风速预报应用研究[J]. *海洋学报*, 2020, 42(1): 67-77.
- WANG G S, WANG X D, HOU M, et al. Research on application of LSTM deep neural network on historical observation data and reanalysis data for sea surface wind speed forecasting[J]. *Acta Oceanologica Sinica*, 2020, 42(1): 67-77.
- [14] 徐凌宇, 张高唯, 江湾湾, 等. 深度学习神经网络及其在海洋环境信息挖掘预测中的应用[J]. *海洋信息*, 2018, 33(1): 17-23.
- XU L Y, ZHANG G W, JIANG W W, et al. Deep learning neural network and its application in marine environmental information mining prediction[J]. *Marine Information*, 2018, 33(1): 17-23.
- [15] 查铖, 贺琪, 宋巍, 等. 结合注意力机制的区域型海表面温度预报算法[J]. *海洋通报*, 2020, 39(2): 191-199.
- ZHA C, HE Q, SONG W, et al. Regional sea surface temperature prediction algorithm combined with attention mechanism[J]. *Marine Science Bulletin*, 2020, 39(2): 191-199.
- [16] 焦艳, 黄菲, 高松, 等. 基于长短时记忆神经网络的辽东湾海冰延伸期预报方法研究[J]. *中国海洋大学学报*, 2020, 50(6): 1-11.
- JIAO Y, HUANG F, GAO S, et al. Research on extended-range forecast model of sea ice in the Liaodong bay based on long short term memory network[J]. *Periodical of Ocean University of China*, 2020, 50(6): 1-11.
- [17] 林琪凡, 耿旭朴, 谢婷, 等. 基于长短期记忆神经网络的西太平洋暖池变化预测[J]. *厦门大学学报(自然科学版)*, 2021, 60(5): 927-936.
- LIN Q F, GENG X P, XIE T, et al. Trend prediction of Western Pacific warm pool based on long short-term memory neural networks[J]. *Journal of Xiamen University (Natural Science)*, 2021, 60(5): 927-936.
- [18] 宣良龙. 海洋水声环境效应建模与应用[M]. 北京: 科学出版社, 2012.
- DA L L. Modeling and application of underwater acoustic environmental effect[M]. Beijing: Science Press, 2012.
- [19] LIU H L, LIN P F, ZHENG W P, et al. A global eddy-resolving ocean forecast system in China - LICOM Forecast System (LFS) [J]. *Journal of Operational Oceanography*, 2023, 16(1): 15-27.

Research on ocean temperature profile forecasts based on Long Short -Term Memory neural network

FAN Peiqin¹, GUO Wuhong¹, TANG Shuai¹, ZHANG Chi¹, QU Hongyue²

(1. Navy Submarine Academy, Qingdao 266199, China; 2. Pilot National Laboratory for Marine Science and Technology, Qingdao 266000, China)

Abstract: Based on ocean temperature historical observations and ocean model data, the short-term forecasting method for ocean temperature profile is studied using Long Short -Term Memory (LSTM) neural network. For the ocean temperature profile forecasts at (17°46.91'N, 112°03.24'E) in the South China Sea, three sample sets of observation, forecasts, mixed observation-forecasts are constructed using the observational and forecasting data. Based on the LSTM neural network, a many-to-many ocean temperature profile forecasting model composed of encoder and decoder is established, and model training and verification are carried out. The results show that the model has a high forecasting accuracy and a good stability for processing small sample problems. Utilizing deep layered forecasting method can effectively improve the forecasting accuracy and generalization ability. The forecasting error of mixed sample set decreases significantly in comparison with that of observation sample set, which provides an idea for marine environment forecasts of small sample problems.

Key words: deep learning; LSTM; ocean temperature; short-term prediction