

BP 和 RBF 神经网络应用于海表温盐短期预测效果对比

李亚蒙^{1,2}, 丁军航^{1,3}, 孙宝楠^{2,4,5,6}, 官 晟^{2,4,5,6*}

(1. 青岛大学 自动化学院, 山东 青岛 266071;

2. 自然资源部 第一海洋研究所, 山东 青岛 266061;

3. 山东省工业控制技术重点实验室, 山东 青岛 266071;

4. 自然资源部 海洋环境科学与数值模拟重点实验室, 山东 青岛 266061;

5. 山东省海洋环境科学与数值模拟重点实验室, 山东 青岛 266061;

6. 青岛海洋科学与技术试点国家实验室 区域海洋动力学与数值模拟功能实验室, 山东 青岛 266237)

摘要: 为了对海表温度 (Sea Surface Temperature, SST) 和海表盐度 (Sea Surface Salinity, SSS) 数据进行精确的短期预报, 基于多站位海洋观测浮标获取的海表温度和海表盐度数据, 利用反向传播 (Back Propagation, BP) 和径向基函数 (Radial Basis Function, RBF) 两种神经网络方法开展了短期预测。首先, 在预测时长固定为 5 d 的情况下, 对比不同训练时长的预测结果的均方误差 (Mean Squared Error, MSE), 进而确定以 20 d 的观测数据作为训练集的预测结果均方误差最小。然后, 以 PAPA 站观测浮标获取的 2009 年 1 月、4 月、7 月和 10 月各月的前 20 d 温盐数据作为训练集, 分别训练 BP 和 RBF 神经网络, 将训练好的 2 种神经网络模型应用于各月第 21 至 25 日的温盐数据预测。结果表明: BP 和 RBF 神经网络均能有效预测海表温盐数据的季节性变化, 但 RBF 神经网络对不同预测时间的整体预测效果优于 BP 神经网络。多站点数据的预测实验进一步验证了 RBF 神经网络模型具有较强的适用性和更高的准确性。RBF 神经网络模型可以作为海表温盐数据短期预报的有力工具。

关键词: BP 神经网络模型; RBF 神经网络模型; 海洋观测浮标; 数据预测; 海表温度; 海表盐度

中图分类号: P731.12; P731.11; TP183

文献标志码: A

文章编号: 1671-6647(2022)02-0220-13

doi: 10.12362/j.issn.1671-6647.2022.02.006

引用格式: 李亚蒙, 丁军航, 孙宝楠, 等. BP 和 RBF 神经网络应用于海表温盐短期预测效果对比[J]. 海洋科学进展, 2022, 40(2): 220-232. LI Y M, DING J H, SUN B N, et al. Comparison of short-term prediction effects of the sea surface temperature and salinity based on BP and RBF neural network[J]. Advances in Marine Science, 2022, 40(2): 220-232.

20 世纪以来, 在全球气候变暖趋势下, 极端气候及灾害事件频发, 海洋环境也变得愈加复杂。要准确掌握海洋环境变化的规律, 进而增强人类认识海洋、开发海洋的能力, 需要具备比以往更准确、更及时和更全面的观测手段。海洋观测浮标可利用其搭载的各类传感器和设备, 测量包括温度和盐度等海洋及大气环境要素, 并实时将观测数据回传岸基台站。海洋观测浮标具有长时间连续、定点观测能力, 搭载传感器种类、数量较多, 其观测方式与海洋环境变化特征相吻合, 是海洋立体观测网络中重要的节点形式。对观测浮标获取数据的有效利用与深度分析, 可为全球气候变化趋势研究、海洋灾害预警和人类海上作业等领域提供有效的数据支撑^[1-3]。

海洋观测浮标获取的多参数数据可以帮助人们了解海洋环境历经的状态, 但对于人类海上活动而言, 更

收稿日期: 2021-11-04

资助项目: 科技部重大科学仪器专项——海洋物性参数监测仪(2018YFF01014105)

作者简介: 李亚蒙(1996—), 女, 硕士研究生, 主要从事智能控制与数据资料预测分析方面研究. E-mail: 516553884@qq.com

* **通信作者:** 官 晟(1972—), 男, 研究员, 博士, 主要从事海洋环境观测技术与智能控制方面研究. E-mail: gsh30@163.com

(王 燕 编辑)

需要从已有数据中预测未来将发生的变化,由此产生了人们对海洋环境要素进行预测的需求。相对于长期预测而言,海洋环境要素的时间序列短期预测面临更多偶然因素干扰,实现精确预测难度更大,但短期预测的结果对海上工程作业在安全保障、方案选择和成本控制等方面具有很高的应用价值。

对于绝大多数由实际观测资料构成的动力系统,往往表现出非线性、非平稳性双重特征。即系统各要素非独立,存在耦合作用,同时观测资料的时间序列又表现出具有趋势或周期性变化特性。对于这些动力系统,通常利用统计预测理论和非线性预测理论开展预测,由于其对信号平稳性的假定,有可能导致预测水平低下^[4]。也有研究^[5]通过对数变换或差分变换将相应的过程平稳化,对能够寻找近似平稳信号分量的非平稳过程取得了不错的预测效果。越来越多的研究^[6-12]将神经网络预报法应用于气象和水文等非线性预测领域。常见的神经网络预报法包括反向传播(Back Propagation, BP)神经网络、径向基函数(Radial Basis Function, RBF)神经网络和差分自回归移动平均模型(Autoregressive Integrated Moving Average model, ARIMA)等。相比之下,海洋动力学预报法需要大型计算机对流体力学微分方程进行长时间积分,所需计算成本和对研究者建模水平的要求都较高。而神经网络方法逼近能力强、收敛速度快,具有信息分布式存储和并行协同处理的特点^[13],实现相对简单,在近年海水温、盐预报应用研究中更为普遍。

基于 BP 神经网络和随机森林算法的评估模型^[14]可以用来精准估算近岸海域的海表盐度。高国栋等^[15]基于 RBF 神经网络模型建立了海水盐度的预测方法,并证明了该模型相较于 BP 神经网络模型收敛速度更快、精度更高。建立基于机器学习的反演模型并利用遥感技术进行大尺度动态模拟^[16],或利用卫星辅助数据建立 BP 神经网络预测模型^[17]均能较为精确地预测海表盐度。王颖超等^[17]利用 BP 神经网络方法预测的海表盐度以及土壤水分和海洋盐度(Soil Moisture and Ocean Salinity, SMOS)卫星提供的 3 个粗糙度模型给出了盐度产品与 ARGO 浮标实测盐度的均方根误差(Root Mean Square Error, RMSE),分别为 0.847、2.041、2.028 和 2.081。基于空间遥感数据,使用神经网络算法预测^[18-19]或者将深度学习神经网络与数值估值器相结合^[20]均能有效预测海表温度。YANG 等^[18]使用长短期记忆(Long Short-Term Memory, LSTM)神经网络预测了渤海 1 d 的海表温度数据,支持向量机、支持向量回归模型、FC-LSTM 等算法的预测结果与遥感实测值的 RMSE 为 0.076~0.399,预测精度为 98.49%~99.58%,证明了所提出的神经网络预测方法可行,但也发现数据集空间分辨率越高,精确预测难度越大。

以往研究结果证明了多种神经网络预测技术能够有效应用于海表温、盐时间序列数据的预测工作。然而在这些研究中,未讨论不同训练时长和预测时长等因素对预测精度的影响,也没有讨论算法的多海域适用性等问题。基于上述情况,本文将利用 BP 和 RBF 两种神经网络方法开展研究,以期获得一种具有稳定预报准确率,并具有较广泛适用性的海表温、盐短期预报方法。

1 数据与方法

1.1 数据来源

本研究数据主要来自美国国家海洋和大气管理局(National Oceanic and Atmospheric Administration, NOAA)组织实施的 OceanSITES 观测系统中 PAPA 站(145°W, 50°N)、Stratus 站(85°W, 20°S)、KEO 站(145°E, 32°N)、WHOTS 站(160°W, 25°N)和 JKEO 站(146°E, 38°N)等多个浮标站位于 2009—2012 年的监测数据。这些数据包括 0~300 m 深度的温度、0~200 m 深度的盐度和密度、表层海流、纬向和经向海流的声学多普勒剖面等,以及风速、气温、相对湿度和大气压等气象观测数据。本文采用 MatLab 软件及其机器学习工具箱进行短期预测方法研究。

1.2 实验方法

1.2.1 BP 神经网络预测模型的构建

BP 神经网络是一种按照误差逆向传播算法训练的多层前馈神经网络,是目前应用最为广泛的神经网络

络。BP 网络按结构分为输入层、隐藏层和输出层,如图 1 所示。

对于隐藏层,利用激活函数处理输入可以得到隐藏层输出,公式为:

$$y_j = f(\text{net}_j), \quad (1)$$

式中, $f(\text{net}_j)$ 为隐藏层激活函数,其中 $\text{net}_j = \sum_{i=1}^m \omega_{ij} x_i$ 。

对于输出层,利用激活函数处理隐藏层输出可以得到神经元输出,公式为:

$$h_k = f(\text{net}_k), \quad (2)$$

式中, $f(\text{net}_k)$ 为输出层激活函数,其中 $\text{net}_k = \sum_{j=1}^n v_{jk} y_j$ 。

BP 神经网络的优点是拥有很好的非线性映射能力和灵活的网络结构。网络各层的神经元个数可根据具体情况设定,并且其性能也会随着结构的差异而有所不同。但是 BP 神经网络也存在学习速度慢、因容易陷入局部极小值而无法得到全局最优值等主要缺陷。建立 BP 神经网络模型的步骤主要有 5 步^[21]。

步骤一:读取研究对象数据集,如海表温度数据或海表盐度数据,对缺损值和突兀点进行预处理。

步骤二:划分训练集和测试集,对海表温度和海表盐度数据进行归一化处理。一般情况下,训练集的数据量会影响 BP 神经网络的预测效果。因此,以 PAPA 站 2009—2012 年每年 4 月海表温度数据为样本,在预测时长固定为 5 d 的情况下,对比 5 d、10 d、15 d、20 d 和 25 d 不同时长训练海表温度数据集预测结果的均方误差,结果如表 1 所示。由表 1 可知,选取 20 d 的观测海表温度数据作为训练集,预测结果相对最优,训练时长也在可接受范围内。

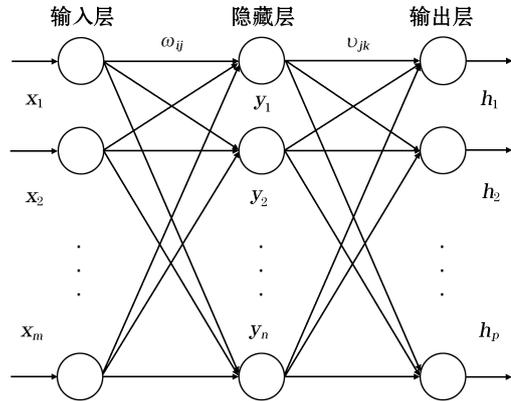
表 1 2009—2012 年每年 4 月 BP 神经网络不同训练时长对应海表温度预测结果的均方误差

Table 1 MSE of SST prediction results of different training time by BP neural network in April 2009—2012

时 间	5 d MSE	10 d MSE	15 d MSE	20 d MSE	25 d MSE
2009 年 4 月	7.15×10^{-2}	3.39×10^{-2}	2.21×10^{-2}	2.07×10^{-2}	2.32×10^{-2}
2010 年 4 月	6.72×10^{-2}	4.67×10^{-2}	2.14×10^{-2}	1.91×10^{-2}	2.24×10^{-2}
2011 年 4 月	7.35×10^{-2}	8.09×10^{-4}	5.81×10^{-4}	5.68×10^{-4}	5.83×10^{-4}
2012 年 4 月	0.56×10^{-2}	0.44×10^{-2}	0.24×10^{-2}	0.22×10^{-2}	0.23×10^{-2}

步骤三:确定输入层节点和输出层节点数。实验中将历史海表温度和海表盐度数据作为输入,因此,选取输入层节点数为 1,分别选择下一时刻的海表温度和海表盐度作为输出,因此,实验中输出层节点数为 1。

步骤四:确定网络各层之间的传递函数和训练算法。BP 神经网络使用最多的激活函数为 sigmoid 函数,其输入值可取任意值,输出值范围为 0~1。本文选择隐藏层激活函数为 tansig,其输入值为任意值,输出值为 -1~1。选择输出层激活函数为 purelin,其输入和



注: ω_{ij} ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$) 为输入层到隐藏层的权值, v_{jk} ($k = 1, 2, \dots, p$) 为隐藏层到输出层的权值, x_i 为神经元输入, y_j 为隐藏层输出, h_k 为神经元输出。

图 1 BP 神经网络结构

Fig.1 Structure diagram of BP neural network

表 2 BP 神经网络隐藏层节点数对应海表温度训练集的均方误差

Table 2 The MSE of SST training set corresponding to the number of hidden layer nodes of BP neural network

隐藏层节点数	训练集 MSE	隐藏层节点数	训练集 MSE
2	1.09×10^{-3}	7	1.09×10^{-3}
3	1.11×10^{-3}	8	1.08×10^{-3}
4	1.47×10^{-3}	9	1.11×10^{-3}
5	1.18×10^{-3}	10	1.09×10^{-3}
6	1.09×10^{-3}	11	1.15×10^{-3}

输出值均为任意值。训练算法为 Levenberg-Marquardt[*trainlm*],该算法作为系统的默认算法,其优势在于针对中等规模的 BP 神经网络,收敛速度最快。

通过网络测试,对比隐藏层节点数对应的训练集均方误差值,可以选择出最优的隐藏层节点个数,并应用于海洋数据预测。同样以 2009—2012 年每年 4 月海表温度的预测为例,设置每月前 20 d 海表温度数据为训练集进行实验,训练集均方误差结果如表 2 所示。由表 2 可知,最优的隐藏层节点数为 8。

步骤五:分析 BP 神经网络的预测值与实测值误差。通过误差分析,研究预测方法的适用性及策略改进。

1.2.2 RBF 神经网络预测模型构建

径向基函数 RBF 神经网络主要用于解决多变量插值的问题。其取值依赖于空间任意一点 x 到中心点 c 的距离,即 $\varphi(x, c) = \varphi \|x - c\|$ 。最常用的径向基函数是高斯核函数^[22],其形式为:

$$R(\|x - x_c\|) = \exp\left(-\frac{\|x - x_c\|^2}{2\sigma^2}\right), \quad (3)$$

式中, x_c 是核函数中心, σ 是函数的宽度参数。

RBF 神经网络属于前向神经网络类型,拥有 3 层前向网络结构,包括输入层、隐藏层和输出层。输入层由信号源结点构成;隐藏层神经单元的激活函数是径向基函数,它是对应中心点径向对称且逐渐衰减的非线性函数, BP 神经网络可以有多个隐藏层,但是 RBF 只有一个隐藏层;输出层对输入模式做出响应。从输入层到隐藏层的变换是非线性的,而从隐藏层到输出层的变换是线性的^[23]。RBF 神经网络结构如图 2 所示。

RBF 神经网络实质上是利用由径向基函数构成的隐藏层,将输入变量直接映射至隐藏空间,通过确定 RBF 中心点来控制映射关系。通常径向基函数对于空间中心点具备径向对称性,当神经元输入距离中心点较远时,神经元的激活程度较低,这属于 RBF 神经网络的局部特性^[24]。隐藏层的作用是将输入向量由低维度 n 映射成高维度 h ,使得向量具有线性可分的特性,通常利用高斯核函数作为激活函数:

$$R(\|x_k - c_i\|) = \exp\left(-\frac{\|x_k - c_i\|^2}{2\sigma^2}\right), \quad (4)$$

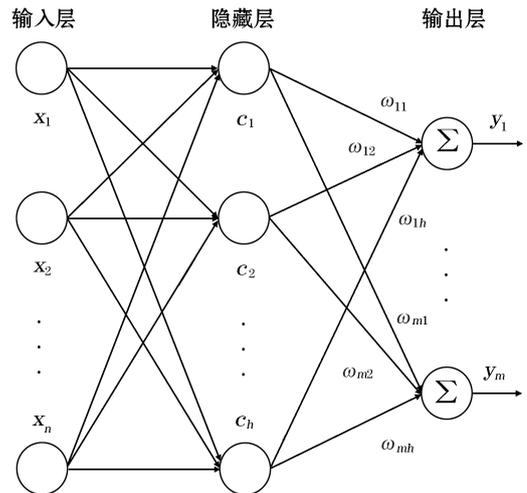
式中, $k = 1, 2, \dots, n; i = 1, 2, \dots, h$ 。

此时网络的输出是隐藏层输出的线性加权和,权即为可调参数,网络输出可表示为:

$$y_j = \sum_{i=1}^h \omega_{ji} \exp\left(-\frac{\|x_k - c_i\|^2}{2\sigma^2}\right), \quad (5)$$

式中, $k = 1, 2, \dots, n, j = 1, 2, \dots, m; \sigma$ 为最小二乘损失函数, $\sigma = \frac{1}{n} \sum_j \|d_j - y_j; c_i\|^2$, 其中 $i = 1, 2, \dots, h; d_j$ 为理想输出值。

RBF 神经网络的设计主要需要确认径向基函数中心、方差和隐藏层到输出层的权值三个参数。对于高斯核函数的径向基网络,首先利用 k -means 聚类得到 h 个中心,然后利用 $\sigma_i = \frac{c_{\max}}{\sqrt{2h}}$ (式中, $i = 1, 2, \dots, h$, c_{\max} 表示选择的中心点之间最大距离)计算径向基函数方差,最后利用最小二乘法计算权值 ω ,计算式为 $\omega =$



注: $x_k (k = 1, 2, \dots, n)$ 为神经元输入, $c_i (i = 1, 2, \dots, h)$ 为隐藏层输出, $\omega_{ji} (j = 1, 2, \dots, m)$ 为隐藏层到输出层权值, y_j 为神经元输出。

图 2 RBF 神经网络结构

Fig.2 Structure diagram of RBF neural network

$\exp\left(\frac{h}{c_{\max}^2} \|x_k - c_i\|^2\right)$ (式中, $k=1, 2, \dots, n; i=1, 2, \dots, h$)。

RBF 模型建立步骤基本与 BP 神经网络类似。同样考虑到训练集的数据量对 RBF 神经网络预测效果的影响,以 2009—2012 年间每年 4 月海表温度数据预测为例,在预测时长固定 5 d 情况下,设置不同训练时长进行对比,预测结果的均方误差如表 3 所示。由表 3 可知,选取 20 d 的海表温度作为训练集,预测结果相对最优。

表 3 RBF 神经网络训练时间对应海表温度预测结果均方误差

Table 3 Mean square error of SST prediction corresponding to different training days by RBF neural network

时 间	5 d MSE	10 d MSE	15 d MSE	20 d MSE	25 d MSE
2009 年 4 月	3.68×10^{-2}	2.82×10^{-2}	1.06×10^{-2}	1.03×10^{-2}	1.04×10^{-2}
2010 年 4 月	1.51×10^{-2}	0.72×10^{-2}	0.39×10^{-2}	0.35×10^{-2}	0.37×10^{-2}
2011 年 4 月	0.71×10^{-2}	6.62×10^{-4}	5.68×10^{-4}	5.61×10^{-4}	5.64×10^{-4}
2012 年 4 月	0.11×10^{-2}	9.08×10^{-4}	9.01×10^{-4}	8.40×10^{-4}	8.97×10^{-4}

1.2.3 评价标准

本文采取均方误差(MSE)和平均绝对误差(Mean Absolute Error, MAE)两个指标来对检测结果进行分析。均方误差是指参数估计值与参数真值之差平方的期望值,可以评价数据的变化程度,MSE 值越小,说明预测模型描述的实验数据具有更好的精确度;平均绝对误差能更好地反映预测值误差的实际情况。2 个指标的计算公式为:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (f_i - y_i)^2, \quad (6)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |(f_i - y_i)|, \quad (7)$$

式中, N 为模型样本数, f_i 为海表温度或海表盐度的实测值, y_i 为海表温度或海表盐度的预测值。

相关系数是一种研究变量之间线性相关程度的量。相关系数以变量与各自平均的差值为基础,按积差的方式进行计算,其表达式为:

$$r = \frac{\sum_{i=1}^N (f_i - \bar{f}_i)(y_i - \bar{y}_i)}{\sqrt{\sum_{i=1}^N (f_i - \bar{f}_i)^2 \sum_{i=1}^N (y_i - \bar{y}_i)^2}}. \quad (8)$$

$|r|$ 处于 0~1 之间, $|r|$ 越接近 1,代表预测值与实测值之间相关程度越高,即预测性能越好; $|r|$ 越接近 0 则代表预测性能越差。

2 结果与分析

海洋表面的海水温度、盐度等物理参数受到波浪、海流、风及海气交换等诸多不稳定因素的影响,有时变化剧烈、随机性较强,利用传统拟合、客观分析等方法难以进行准确的短期预测。

本文首先选择 PAPA 站在 2009 年的 1 月、4 月、7 月和 10 月的海表温度和海表盐度数据作为研究数据集。1 月、4 月、7 月和 10 月是常用的北半球冬季、春季、夏季、秋季四个季节典型环境特点代表月份。PAPA 站浮标搭载传感器对海表温度及海表盐度采样周期均为 1 h。根据前文所述,预测模型所用海表温度及海表盐度实测数据的训练集时间长度为 20 d,短期预测时间长度为 5 d,共 25 d、合计 600 h 的海表温度及海表

盐度数据变化曲线如图 3 和图 4 所示。

从图 3 和图 4 可见,海表温度数据表现出了明显的季节变化特征与日周期的变化特点,在这些趋势性和固定的周期变化基础上,叠加了非线性、随机的扰动。其中 1 月和 10 月海表温度呈明显下降趋势,分别从 5.9 °C 下降到 5.2 °C、从 11.0 °C 下降到 9.0 °C; 4 月和 7 月海表温度呈明显上升趋势,分别从 7.0 °C 上升到 9.0 °C、从 12.0 °C 上升到 13.2 °C。而海表盐度数据的季节性趋势相对不明显,整体变化范围较小,其中 1 月和 4 月大部分时间的海表盐度实测数据均超过 32.50,而 7 月和 10 月大部分时间的海表盐度实测数据却均低于 32.50。

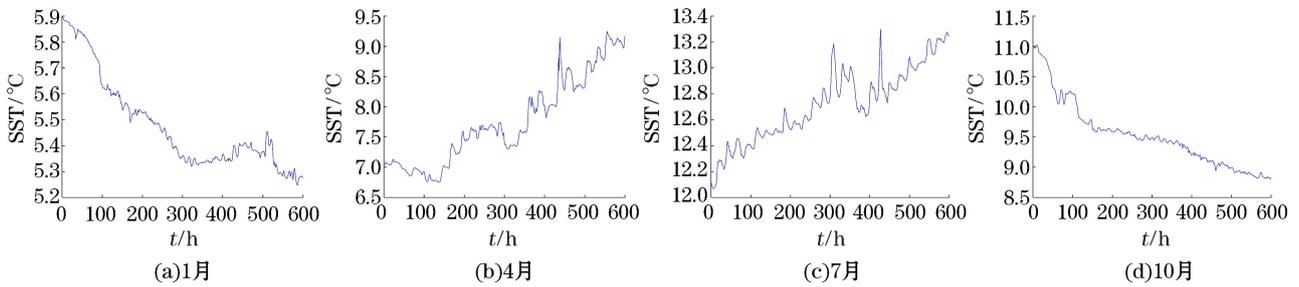


图 3 冬、春、夏、秋四个季节海表温度随时间的变化

Fig.3 Changes of SST with time in winter, spring, summer and autumn

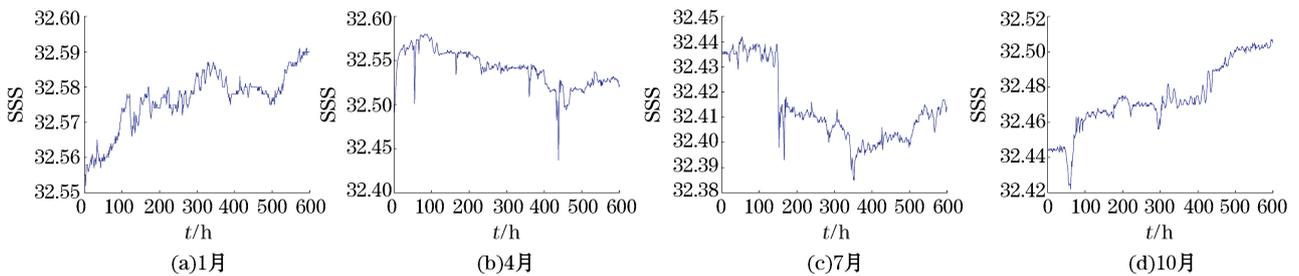


图 4 冬、春、夏、秋四个季节海表盐度随时间的变化

Fig.4 Changes of SSS with time in winter, spring, summer and autumn

2.1 预测对比实验

分别采用 BP 和 RBF 神经网络模型,对 PAPA 站 4 个代表月份的 21 至 25 日、共 120 h 的海表温度与海表盐度进行了短期预报。预报结果如图 5 和图 6 所示。训练集均为各月前 20 d 的实测数据。BP 神经网络隐藏层节点数为 8,RBF 神经网络隐藏层节点数为 8。

从图 5 和图 6 看出,2 种神经网络所获得的预测结果能够反映实测数据的变化趋势,对局部细节也有所反应。2 种方法的预测效果随着预测时长增加均出现相对实测数据的偏离,其中 BP 神经网络的预测结果偏离更为显著。对 2 种方法预测结果的量化评价,主要从均方误差(MSE)和平均绝对误差(MAE)来衡量。

在温度短期预报方面(表 4),2 种方法的 MAE 均远小于 0.5 °C。其中误差超过《海洋调查规范》^[25]规定的 0.02 °C 标准的情况,RBF 预测结果有 1 次,BP 预测结果有 2 次。由于盐度实测值(表 4)较为稳定,各项误差指数均优于温度预报结果。其中误差超过《海洋调查规范》^[25]规定的 0.02 标准的情况,只有 BP 预测结果有 1 次。RBF 对海温和盐度的预测最小 MAE 分别为 0.99×10^{-2} °C 和 9.07×10^{-4} ,BP 对海温和盐度的预测最小 MAE 分别为 1.50×10^{-2} °C 和 0.12×10^{-2} 。

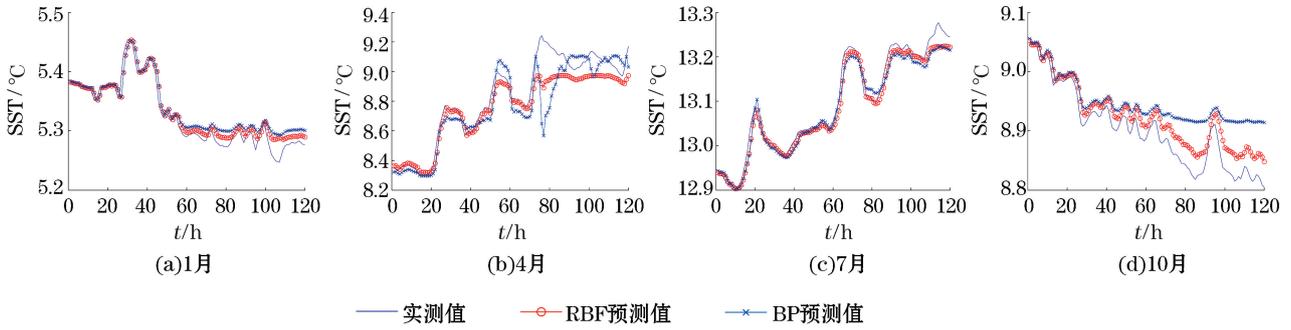


图 5 4 个季节代表月份 21 至 25 日海表温度预测结果

Fig.5 Prediction results of SST from 21 to 25 in representative months of four seasons

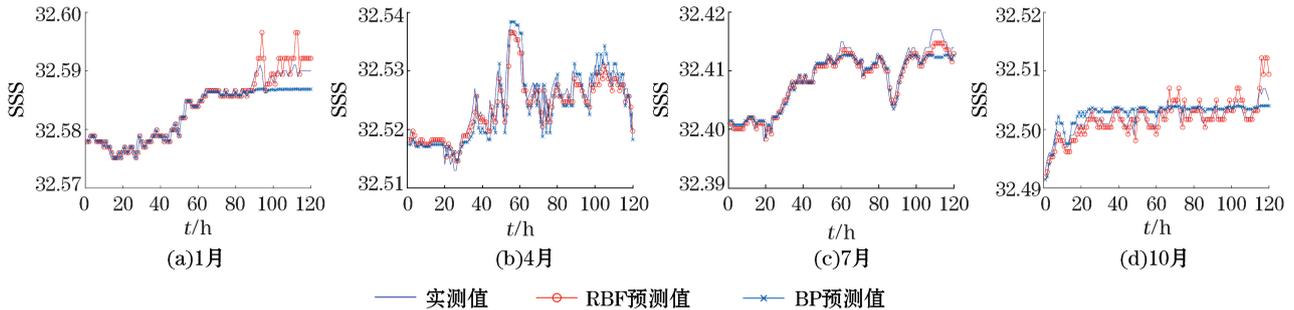


图 6 4 个季节代表月份 21 至 25 日海表盐度预测结果

Fig.6 Prediction results of SSS from 21 to 25 in representative months of four seasons

表 4 2 种预测方法预测海表温度和海表盐度误差分析

Table 4 Error analysis of two methods for predicting SST and SSS

预测方法	1 月海表温度		4 月海表温度		7 月海表温度		10 月海表温度	
	MSE	MAE/°C	MSE	MAE/°C	MSE	MAE/°C	MSE	MAE/°C
RBF	1.82×10^{-4}	0.99×10^{-2}	1.03×10^{-2}	7.12×10^{-2}	3.21×10^{-4}	1.34×10^{-2}	7.15×10^{-4}	2.31×10^{-2}
BP	3.77×10^{-4}	1.50×10^{-2}	2.07×10^{-2}	8.64×10^{-2}	4.79×10^{-4}	1.76×10^{-2}	0.32×10^{-2}	4.50×10^{-2}

预测方法	1 月海表盐度		4 月海表盐度		7 月海表盐度		10 月海表盐度	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
RBF	2.64×10^{-6}	0.11×10^{-2}	5.70×10^{-6}	0.17×10^{-2}	1.45×10^{-6}	9.07×10^{-4}	3.03×10^{-6}	0.12×10^{-2}
BP	3.39×10^{-6}	0.13×10^{-2}	8.98×10^{-6}	0.23×10^{-2}	2.78×10^{-6}	0.12×10^{-2}	5.96×10^{-6}	0.22×10^{-2}

利用式(8)计算 RBF 和 BP 两种方法的预测结果与实测结果的相关系数,结果如表 5 和表 6 所示。从相关系数指标看,RBF 预测方法对海表温度和海表盐度的预测结果,在所有月份均优于 BP 预测方法的结果。受篇幅所限,本文只展示了 2 种预测方法的温度(图 7a)和盐度(图 7b)单月预测结果的相关性趋势,其他 3 个月情况没有大的差异。在相关性趋势图中,RBF 预测方法结果分布点也更集中于 $y = x$ 直线,没有明显的偏离点。因此,就整体预测效果而言,RBF 神经网络方法更精准。

表 5 2 种预测方法的海表温度预测值与实测值相关系数

Table 5 Correlation coefficients between predicted and measured SST of the two methods

预测方法	1 月	4 月	7 月	10 月
RBF	0.98	0.98	0.99	0.98
BP	0.97	0.94	0.98	0.91

表 6 2 种预测方法的海表盐度预测值与实测值相关系数

Table 6 Correlation coefficients between predicted and measured SSS of the two methods

预测方法	1 月	4 月	7 月	10 月
RBF	0.96	0.91	0.98	0.87
BP	0.96	0.89	0.96	0.84

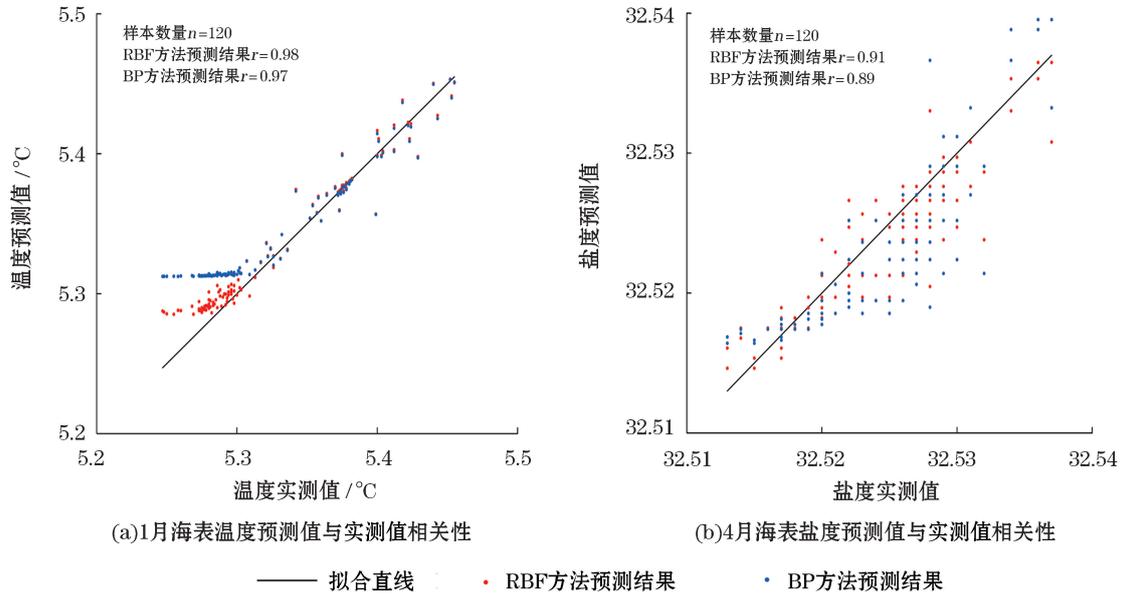


图 7 RBF 和 BP 两种神经网络方法 1 月海表温度、4 月海表盐度预测值与实测值的相关性

Fig.7 Correlation between the predicted and measured values of monthly SST or SSS by the two methods

此外,对于不同季节的海表温度和海表盐度数据,2 种神经网络方法预测效果的准确性均达到海表温度模式预报 0.5 °C 的精度,海表温度和海表盐度的短期预报精度均接近《海洋调查规范》^[25] 规定的 0.02 °C 和 0.02 标准的要求。预测结果的误差指标证明,这 2 种神经网络预测方法具备针对不同季节的海洋环境的适应能力,可以预测整年的海洋温盐数据。

5 d 是典型的短期预报时长,但是在以 20 d 实测数据为训练集的前提下,对 RBF 和 BP 两种神经网络预测方法而言,对能获取更高预测准确率的最佳预测周期可开展进一步研究。

设置预测时长为 1 d、2 d、3 d、4 d 和 5 d,利用 2 种神经网络模型对 PAPA 站 2009 年各季节海表温度和海表盐度进行预测对比,训练集仍选最佳训练时长 20 d。预测值各月 MAE 的结果如图 8 和图 9 所示。由图 8 和图 9 可知,随着测试集预测时长的变化,利用 RBF 神经网络方法预测的海表温度和海表盐度的 MAE 始终低于 BP 神经网络。由于海表温度的变化幅度大于海表盐度,因此随着预测时长的增加,海表温度预测的误差增加较为明显,除个别情形(如 RBF 对 7 月海表温度的预测)外,2 种预测模型的海表温度预测最佳周期可选为 1 d;由于海表盐度的变化幅度较小,因此预测误差随预测时长的变化不太明显,整体趋势为随着预测时长增加,预测误差也对应增大,最佳预测周期可选 2~3 d。在实际应用中,需综合考虑预测成本与准确性的实际需求确定预测时长。

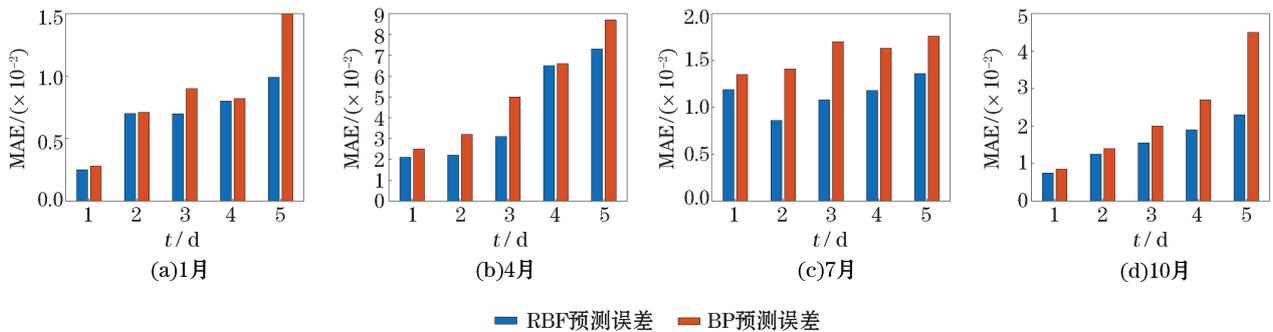


图 8 RBF 和 BP 两种神经网络预测方法在 4 个代表月份以不同预测时长预测海表温度结果的平均绝对误差

Fig.8 The MAE of SST predicted by two prediction methods with different prediction days in four representative months

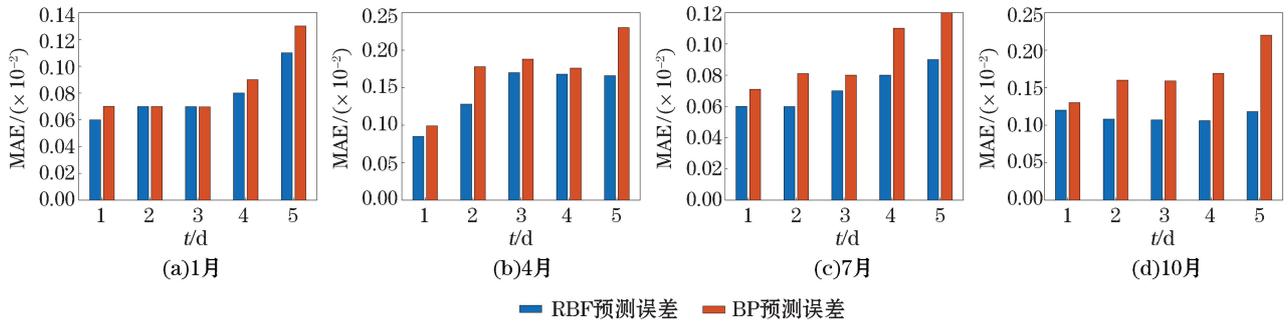


图9 RBF和BP两种神经网络预测方法在4个代表月份以不同预测时长预测海表盐度结果的平均绝对误差

Fig.9 The MAE of SSS predicted by two prediction methods with different prediction days in four representative months

2.2 多站点预测验证

为了验证预测模型的适用性和普遍性,选择PAPA站以外的其他4个站点的浮标数据,利用RBF神经网络模型对海表温盐数据进行预测验证。这4个站点分别为Stratus站($85^{\circ}\text{W}, 20^{\circ}\text{S}$)、KEO站($145^{\circ}\text{E}, 32^{\circ}\text{N}$)、WHOTS站($160^{\circ}\text{W}, 25^{\circ}\text{N}$)和JKEO站($146^{\circ}\text{E}, 38^{\circ}\text{N}$)。各站点海表温度和海表盐度数据采样周期也为1 h,模型参数通过上述实验方法选定,海表温度预测时长为1 d,海表盐度预测时间为3 d。其中,利用位于东太平洋距离较近的KEO站和JKEO站对2010年的海表温度和海表盐度数据进行重复性实验,验证RBF神经网络模型预测能力的可重复性;利用中太平洋相距较远的WHOTS站和西太平洋Stratus站2012年的海表温度和海表盐度数据验证预测模型的普适性。由于数据较多,本文只给出Stratus站($85^{\circ}\text{W}, 20^{\circ}\text{S}$)各季节的预测结果(图10和图11)。各站点的预测误差及相关系数如表7~表10所示。

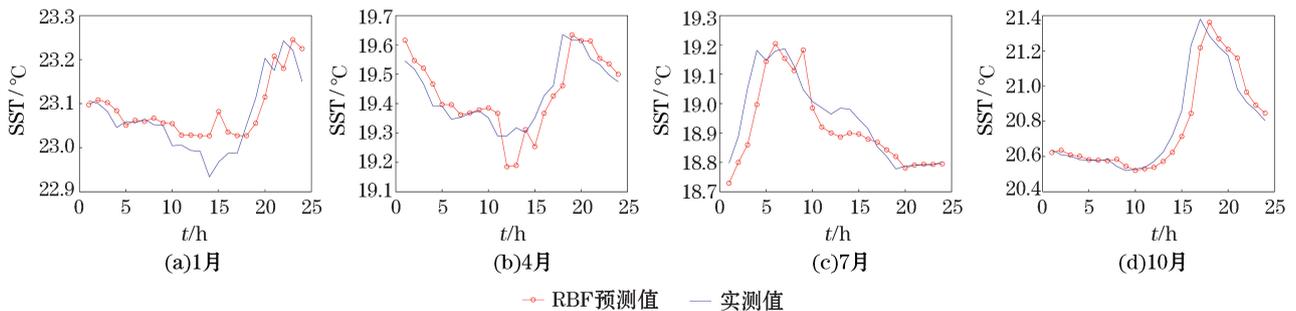


图10 Stratus站4个代表月份1 d海表温度预测结果

Fig.10 1-day SST prediction results of four representative months at Stratus Station

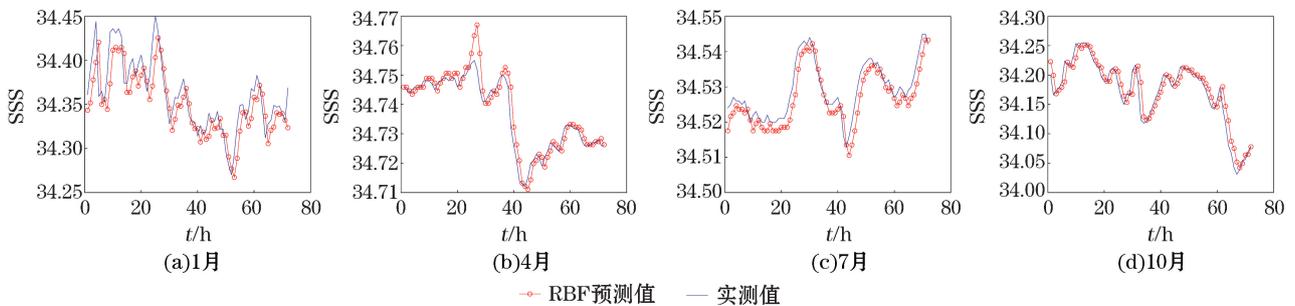


图11 Stratus站4个代表月份3 d海表盐度预测结果

Fig.11 3-day SSS prediction results of four representative months at Stratus Station

表 7 RBF 神经网络海表温度预测误差

Table 7 Prediction error of SST by RBF neural network

站 点	1 月海表温度		4 月海表温度		7 月海表温度		10 月海表温度	
	MSE	MAE/°C	MSE	MAE/°C	MSE	MAE/°C	MSE	MAE/°C
Stratus 站	0.23×10^{-2}	3.82×10^{-2}	0.43×10^{-2}	4.93×10^{-2}	0.57×10^{-2}	5.27×10^{-2}	1.09×10^{-2}	6.28×10^{-2}
KEO 站	7.10×10^{-4}	1.82×10^{-2}	8.14×10^{-2}	1.71×10^{-1}	0.19×10^{-2}	3.67×10^{-2}	0.17×10^{-2}	3.61×10^{-2}
WHOTS 站	9.16×10^{-4}	2.19×10^{-2}	4.29×10^{-4}	1.90×10^{-2}	0.28×10^{-2}	3.32×10^{-2}	0.66×10^{-2}	4.40×10^{-2}
JKEO 站	0.14×10^{-2}	3.14×10^{-2}	0.45×10^{-2}	5.36×10^{-2}	1.00×10^{-2}	7.96×10^{-2}	0.38×10^{-2}	4.93×10^{-2}

表 8 RBF 神经网络海表盐度预测误差

Table 8 Prediction error of SSS by RBF neural network

站 点	1 月海表盐度		4 月海表盐度		7 月海表盐度		10 月海表盐度	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Stratus 站	5.62×10^{-4}	1.91×10^{-2}	1.48×10^{-5}	0.25×10^{-2}	1.39×10^{-5}	0.31×10^{-2}	3.07×10^{-4}	1.28×10^{-2}
KEO 站	8.33×10^{-6}	0.20×10^{-2}	5.14×10^{-5}	0.44×10^{-2}	7.83×10^{-5}	0.60×10^{-2}	9.83×10^{-6}	0.54×10^{-2}
WHOTS 站	9.28×10^{-6}	0.17×10^{-2}	1.81×10^{-5}	0.30×10^{-2}	2.81×10^{-5}	0.38×10^{-2}	3.59×10^{-5}	0.32×10^{-2}
JKEO 站	3.77×10^{-5}	0.32×10^{-2}	1.15×10^{-5}	0.22×10^{-2}	1.02×10^{-4}	0.48×10^{-2}	3.71×10^{-4}	1.46×10^{-2}

表 9 RBF 神经网络预测海表温度的预测值与实测值相关系数

Table 9 Correlation coefficients between predicted and measured SST predicted by RBF neural network

站 点	1 月	4 月	7 月	10 月
Stratus 站	0.83	0.86	0.88	0.93
KEO 站	0.83	0.81	0.96	0.98
WHOTS 站	0.97	0.95	0.99	0.92
JKEO 站	0.98	0.96	0.88	0.92

表 10 RBF 神经网络预测海表盐度的预测值与实测值相关系数

Table 10 Correlation coefficients between predicted and measured SSS predicted by RBF neural network

站 点	1 月	4 月	7 月	10 月
Stratus 站	0.85	0.96	0.94	0.95
KEO 站	0.94	0.90	0.87	0.85
WHOTS 站	0.97	0.95	0.96	0.95
JKEO 站	0.98	0.99	0.93	0.98

由多站点预测实验结果的误差分析和相关系数比对可知,RBF 神经网络模型对于海洋表面温盐数据的超短期预测具备普遍适用性,并能精确拟合出多个海域及其四季海表温盐数据的变化规律。这表明,尽管海表温盐等数据受诸多不稳定因素影响而具备随机性、非线性的特点,但使用 RBF 神经网络能够进行有效、高精度的预测。相较于已往研究成果,应用本方法对独立站位海表温度数据进行 1 d 预测、对海表盐度数据进行 3 d 预测,预测结果在均方误差和平均绝对误差指标方面,均优于已有的海表温度和海表盐度遥感数据预测研究所取得的结果。

3 结 论

基于海洋环境复杂多变,难以做到精准预测的现状,本文利用 BP 和 RBF 神经网络模型对海表温度和海表盐度进行预测,在 MatLab 环境下搭建了 RBF 和 BP 两种神经网络的预测模型,开展了海表温度和海表盐度的短期预测和结果比对研究。主要结论如下:

首先,通过单站点四个季节代表月份的海表温度和海表盐度预测实验,证明了 BP 和 RBF 神经网络模型均能有效模拟海表温度和海表盐度数据的短期变化趋势,对季节性变化均有较强的适应性,RBF 神经网络拥有更高的预测准确性。

其次,基于不同预测时间下的预测值绝对平均误差的对比分析,得出海表温度和海表盐度的最佳预测周期分别为 1 d 和 3 d。同时证明了 RBF 神经网络针对不同预测周期,预测效果始终优于 BP 神经网络。

最后,本文提出并配置各项训练参数的 RBF 神经网络模型,在多站点的海表温度和海表盐度预测实验中,表现出在多个不同的海域,对海表温度和海表盐度普遍适用的精确预测能力。

后续将进一步开展年际预测精度变化与全球变化之间关联性研究,以及对海洋数据预测领域中更多海洋数据如辐照度、风向等进行预测。

参考文献(References):

- [1] 王宁. 面向机器学习的海洋环境数据分析与预测研究[D]. 唐山: 华北理工大学, 2020. WANG N. Research of analysis and prediction for marine environment data based on machine learning[D]. Tangshan: North China University of Science and Technology, 2020.
- [2] 王军成. 海洋资料浮标原理与工程[M]. 北京: 海洋出版社, 2013: 3-7. WANG J C. Principle and engineering of ocean data buoy[M]. Beijing: China Ocean Press, 2013: 3-7.
- [3] 王波, 李民, 刘世萱, 等. 海洋资料浮标观测技术应用现状及发展趋势[J]. 仪器仪表学报, 2014, 35(11): 2401-2413. WANG B, LI M, LIU S X, et al. Current status and trend of ocean data buoy observation technology application[J]. Chinese Journal of Scientific Instrument, 2014, 35(11): 2401-2413.
- [4] 王革丽, 杨培才, 毛宇清. 基于支持向量机方法对非平稳时间序列的预测[J]. 物理学报, 2008(2): 714-719. WANG G L, YANG P C, MAO Y Q. On the application of non-stationary time series prediction based on SVM method[J]. Acta Physica Sinica, 2008(2): 714-719.
- [5] 王革丽, 杨培才, 卞建春, 等. 一个包含外强迫因子的非平稳时间序列的预测方法[J]. 科学通报, 2011, 56(25): 2107-2110. WANG G L, YANG P C, BIAN J C, et al. A novel approach in predicting non-stationary time series by combining external forces[J]. Chinese Science Bulletin, 2011, 56(25): 2107-2110.
- [6] 余开华. 小波神经网络模型在河道流量水位预测中的应用[J]. 水资源与水工程学报, 2013, 24(2): 204-208. YU K H. Application of RBF and GRNN neural network model in forecast of water runoff and head[J]. Journal of Water Resources and Water Engineering, 2013, 24(2): 204-208.
- [7] 赵世新, 张晨, 李静怡, 等. 基于 BP 神经网络方法的引滦入津工程黎河段水质预测研究[J]. 安全与环境学报, 2010(2): 93-96. ZHAO S X, ZHANG C, LI J Y, et al. Predicted water-quality evaluation of Lihe section in Luanhe diversion to Tianjin water project based on BP neural network[J]. Journal of Safety and Environment, 2010(2): 93-96.
- [8] 董世超. 基于 ARIMA-BP 神经网络模型海流流速预测研究[J]. 中国科技信息, 2014(2): 86-88. DONG S C. Current prediction research based on ARIMA-BP neural network abstract[J]. China Science and Technology Information, 2014(2): 86-88.
- [9] 吴晓玲, 王船海. 水动力学预报模型中系统噪声均值的空间分布[J]. 人民长江, 2008(5): 15-16, 22. WU X L, WANG C H. Spatial distribution of system noise mean in hydrodynamic prediction model[J]. Yangtze River, 2008(5): 15-16, 22.
- [10] ALAVI A H, GANDOMI A H, GANDOMI M, et al. Prediction of maximum dry density and optimum moisture content of stabilised soil using RBF neural networks[J]. The IES Journal Part A: Civil & Structural Engineering, 2009, 2(2): 98-106.
- [11] 郭鑫, 韩震, 张雪薇, 等. HY-2B 卫星散射计神经网络多区间风速反演[J]. 海洋科学进展, 2021, 39(2): 268-278. GUO X, HAN Z, ZHANG X W, et al. Multi interval wind speed inversion for HY-2B satellite scatterometer based on neural network[J]. Advances in Marine Science, 2021, 39(2): 268-278.
- [12] 杜星, 孙永福, 宋玉鹏, 等. 基于人工神经网络的海底细粒土工程分类研究[J]. 海洋科学进展, 2021, 39(1): 110-117. DU X, SUN Y F, SONG Y P, et al. Study on engineering classification of seafloor fine grained soil based on artificial neural network[J]. Advances in Marine Science, 2021, 39(1): 110-117.
- [13] 胡方明, 简琴, 张秀君. 基于 BP 神经网络的车型分类器[J]. 西安电子科技大学学报, 2005(3): 439-442. HU F M, JIAN Q, ZHANG X J. The classifier of car types using BP neural networks[J]. Journal of Xidian University, 2005(3): 439-442.
- [14] LIU M L, LIU X N, JIANG J L, et al. Artificial neural network and random forest approaches for modeling of sea surface salinity[J]. International Journal of Remote Sensing Applications, 2013, 3(4): 229-234.
- [15] 高国栋, 张文孝, 慕光宇. RBF 网络和 BP 网络在海水盐度建模中的比较研究[J]. 海洋通报, 2011, 30(1): 12-15. GAO G D, ZHANG W X, MU G Y. A comparative study on RBF network and BP network in the model of salinity[J]. Marine Science Bulletin, 2011, 30

- (1): 12-15.
- [16] 徐朝. 基于机器学习的海面盐度遥感反演模型[D]. 北京: 中国地质大学(北京), 2016. XU Z. Remote sensing inversion models of sea surface salinity based on machine learning method[D]. Beijing: China University of Geosciences (Beijing), 2016.
- [17] 王颖超, 柳青青, 李洪平, 等. 海表面盐度的高精度预测模型[J]. 海洋科学进展, 2021, 39(1): 37-44. WANG Y C, LIU Q Q, LI H P, et al. High-precision prediction model for sea surface salinity[J]. Advances in Marine Science, 2021, 39(1): 37-44.
- [18] YANG Y T, DONG J Y, SUN X, et al. A CFCC-LSTM model for sea surface temperature prediction[J]. IEEE Geoscience, Remote Sensing Letters, 2018, 15(2): 207-211.
- [19] 胡泽煜. 基于非平稳时间序列的海面温度预测研究[D]. 上海: 上海海洋大学, 2021. HU Z Y. Research on sea surface temperature prediction based on non-stationary time series[D]. Shanghai: Shanghai Ocean University, 2021.
- [20] SARKAR P P, JANARDHAN P, ROY P. Prediction of sea surface temperatures using deep learning neural networks[J]. SN Applied Sciences, 2020, 2(8): 1458.
- [21] 张秀菊, 王柳林, 李秀平, 等. 基于 BP 神经网络的潇河流域水质预测[J]. 水资源与水工程学报, 2021, 32(5): 19-26. ZHANG X J, WANG L L, LI X P, et al. Water quality prediction of the Xiaohe river basin based on BP neural network model[J]. Journal of Water Resources and Water Engineering, 2021, 32(5): 19-26.
- [22] 赵健, 樊彦国, 丁宁. 基于最小二乘与径向基函数神经网络的海平面变化预测[J]. 海洋科学, 2018, 42(5): 92-97. ZHAO J, FAN Y G, DING N. Sea level anomaly forecasting using least square and the radial basis function neural network[J]. Marine Science, 2018, 42(5): 92-97.
- [23] CHEN S, COWAN C F, GRANT P M. Orthogonal least squares learning algorithm for radial basis function networks[J]. IEEE Transaction on Neural Networks, 1991, 2(2): 302-309.
- [24] PARK J, SANDBERG I W. Universal approximation using radial-basis-function networks[J]. Neural Computation, 2014, 3(2): 246-257.
- [25] 国家海洋标准计量中心. 海洋调查规范: 第 2 部分 海洋水文观测: GB/T 12763.2—2007[S]. 北京: 中国标准出版社, 2007: 5-8. National Center of Ocean Standards and Metrology. Specification for oceanographic survey: part 2 marine hydrological observation: GB/T 12763.2—2007[S]. Beijing: Standards Press of China, 2007: 5-8.

Comparison of Short-Term Prediction Effects of the Sea Surface Temperature and Salinity Based on BP and RBF Neural Network

LI Ya-meng^{1,2}, DING Jun-hang^{1,3}, SUN Bao-nan^{2,4,5,6}, GUAN Sheng^{2,4,5,6}

(1. School of Automation, Qingdao University, Qingdao 266071, China;

2. First Institute of Oceanography, MNR, Qingdao 266061, China;

3. Shandong Key Laboratory of Industrial Control Technology, Qingdao 266071, China;

4. Key Laboratory of Marine Science and Numerical Modeling, MNR, Qingdao 266061, China;

5. Shandong Key Laboratory of Marine Science and Numerical Modeling, Qingdao 266061, China;

6. Laboratory of Regional Oceanography and Numerical Modeling,

Pilot National Laboratory of Marine Science and Technology, Qingdao 266237, China)

Abstract: In order to make an accurate short-term prediction of Sea Surface Temperature (SST) and Sea Surface Salinity (SSS) data, the short-term prediction results of Back Propagation (BP) and Radial Basis Function (RBF) neural network methods are compared and analyzed by using the SST and SSS data obtained by multi-station marine observation buoys. Firstly, when the prediction days are fixed to 5 days, the Mean Squared Errors (MSE) of the prediction results of different training days are compared, and then determine that the MSE with 20 days of observation data as the training set is the smallest. Then, taking the SST and SSS data of the first 20 days in January, April, July and October 2009 obtained by the observation buoy of PAPA station as the training set, BP and RBF neural networks are trained respectively. There-

after, the trained two neural network models are applied to the prediction of SST and SSS data from the 21st to 25th days of each month. The results show that both BP and RBF neural network can effectively predict the seasonal changes of SST and SSS data. But the comparative experiments of different prediction days show that the overall prediction effect of RBF neural network is better than BP neural network. Finally, through the prediction experiment of multisite data, it is verified that the RBF neural network model has strong applicability and higher accuracy, and can become a powerful tool for short-term prediction of SST and SSS data.

Key words: BP neural network model; RBF neural network model; marine observation buoys; data prediction; sea surface temperature; sea surface salinity

Received: November 4, 2021