

基于支持向量机预测安达曼海南部内孤立波传播特性

陆可潇, 王 晶, 魏 鑫

(中国海洋大学 物理与光电工程学院, 山东 青岛 266100)

摘要: 内孤立波是发生在密度稳定层化海水中的—种特殊的海洋内波。预测内孤立波传播难度较大。本文提出了一种方法, 利用美国麻省理工学院大气环流模型(MITgcm)的内孤立波模型计算了大量模拟数据, 建立数据库。采用机器学习的方法, 建立一个基于支持向量机(support vector machine, SVM)的安达曼海南部内孤立波传播预测模型。最后运用安达曼海南部的 Sentinel-1A 合成孔径雷达(SAR)图像对内孤立波传播预测模型结果进行检验。结果表明: 基于 SVM 的内孤立波传播时间预测模型预测的时间平均绝对百分比误差为 8.43%, 平均绝对误差为 1.00 h。基于 SVM 的内孤立波到达位置预测模型预测的位置平均绝对百分比误差为 0.071%, 平均绝对误差为 0.069°。基于 SVM 的内孤立波振幅预测模型预测的振幅范围为 23.80~84.98 m。

关键词: 内孤立波; MITgcm 数值模拟; 支持向量机(SVM); 预测

中图分类号: TP79 **文献标识码:** A **文章编号:** 1000-3096(2021)05-0031-08

DOI: 10.11759/hyxx20201105007

海洋内波是发生在密度稳定层化的海水内部的一种波动, 其最大振幅出现在海洋内部^[1]。内孤立波是一种具有大振幅、单向传播、传播速度快等特点的非线性内波^[2]。内孤立波具有很强的垂向速度, 在海水营养物质的混合输送中起着重要作用^[3]。由于内孤立波在传播过程中伴随着巨大的扰动, 所以对海洋石油钻井平台和潜艇航行有很大的影响。因此, 内孤立波的预测工作具有重要的现实意义。印度洋的安达曼海几十年来一直是研究内孤立波的经典区域。安达曼海的内孤立波最早是在油气勘探中发现的。1975 年底和 1976 年初, 美国埃克森石油公司在安达曼海南部实施了为期 4 个月的测量计划和随后的钻井作业, 初步的测量确定了该区域存在大的内波, 观测到的内波流速高达 1.8 m/s。1976 年 10 月, 在安达曼海南部的一个 4 d 测量计划中也获得了内波数据。Osborne 等^[4]根据现场观测内波数据证实了内孤立波的存在。在安达曼海北部一名海员记录的一次特殊事件中, 内孤立波通过时, 观察到一个石油钻井平台倾斜了大约 3°^[5]。

人们对安达曼海的内孤立波进行了系统研究。Alpers 等^[6]利用 SAR 图像研究了安达曼海内孤立波的生成位置与大陆架的相互作用。Hyder 等^[5]在安达曼海用实测的潮汐来预测内孤立波的发生。安达曼群岛东北的内孤立波只出现在大潮中, 当潮差超过

1.5 m 时, 内孤立波发生的概率随潮差的增大而增大。Jackson^[7]利用中分辨率成像光谱仪(MODIS)观测, 绘制了一张内孤立波的全球分布图, 并证明内孤立波在安达曼海非常活跃。Da Silva 等^[8]利用 SAR 图像研究了安达曼海的第一模态与第二模态内孤立波的潜在生成区域与机制。Shimizu 等^[9]使用高分辨率麻省理工学院通用环流模型 MITgcm 模拟和扩展 Miles 理论研究了安达曼海内孤立波斜相互作用。Sun 等^[10]利用 MODIS 影像给出了安达曼海内波的时空分布、产生源和传播特征。Raju 等^[11]利用 MODIS 和 SAR 图像来探测安达曼海内孤立波的潜在生成地点及其传播特性。

综上所述, 安达曼海内孤立波的研究大多集中在利用卫星图像来观测内孤立波的生成位置和时空分布, 而对于内孤立波的 MITgcm 数值模拟和预测研究非常少。安达曼海内孤立波多发, 可分为南部、中部和北部 3 个区域。因为这 3 个区域内孤立波大

收稿日期: 2020-11-05; 修回日期: 2021-01-07

基金项目: 国家重点研发项目(2017YFC1405600); 国家自然科学基金项目(61871353)

[Foundation: National Key R&D Program of China, No. 2017YFC1405600; Natural Science Foundation of China, No. 61871353]

作者简介: 陆可潇(1995—), 男, 山东青岛人, 硕士研究生, 研究方向: 海洋遥感, E-mail: lukexiao1234@163.com; 王晶(1962—), 通信作者, 教授, 博士生导师, 主要从事光纤光学和海洋遥感研究, E-mail: wjing@ouc.edu.cn

多数是由半日潮与海山相互作用产生，机理相同，因此只选南部为研究对象，作为内孤立波传播预测的示范区域。运用 MITgcm 模式来计算内孤立波的生成和传播，提取大量的内孤立波数据建立数据库，建立基于支持向量机(support vector machine, SVM)的安达曼海南部内孤立波传播预测模型，利用 SAR 图像检验结果表明模型是有效的。

1 数据来源与数据库建立

本文使用二维 MITgcm 模式来模拟安达曼海南部内孤立波的生成和传播，并从中提取数据建立数据库。模拟海域选取在安达曼海南部(6.08°N,

94.6°E~97.1°E)。MITgcm 采用非静力近似模拟条件，内部区域的水平网格分辨率为 300 m(纬向)，垂直方向上的最大水深为 1 448 m，网格分为 42 层，分辨率从 5 m(海表面)至 584 m(海底)依次递增。为避免模式内生成的内孤立波及其他波动传到边界处发生反射，MITgcm 采用海绵边界条件^[12]。将安达曼海的地形、温度、盐度和潮汐等数据输入到 MITgcm 模式中，地形数据取自 ETOPO1 Global Relief Model (2009)，地形剖面图如图 1 所示。温度和盐度数据均来自于 World Ocean Atlas(2013)的年平均数据，潮汐数据来自美国俄勒冈大学开发的全球潮汐模式 TPXO7.2 的反演结果。

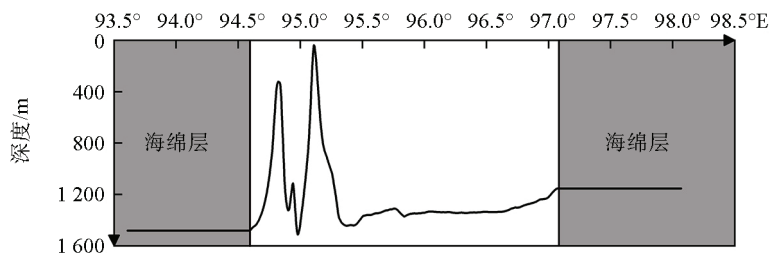


图 1 地形剖面图

Fig. 1 Terrain profiles

安达曼海内孤立波生成被广泛认为是潮地相互作用机制。通过全球潮汐模式 TPXO7.2 获得安达曼海 13 个分潮的速度纬向分量，如表 1 所示。从表 1 中可以看出半日频率的正压潮在此地区占主导地位。因此采用 4 个半日潮(M_2 、 S_2 、 N_2 、 K_2)驱动 MITgcm 模型。把每天的潮汐速度作为初始强迫场输入到 MITgcm 模型中，得到每个潮汐下的内孤立波的产生和传播过程。本文运用 MITgcm 模拟计算了潮汐速度 0.05~0.1 m/s 时的内孤立波生成与传播，提取了 7 584 条内孤立波的到达时间、位置和振幅。建立了潮汐速度振幅、经度、纬度、水深、时间和振幅 6 个参量建立数据库，把数据库的数据进行随机排序，如表 2 所示。运用数据库中不同的参量建立不同的模型，来预测内孤立波的传播时间、到达位置和振幅。

表 1 TPXO7.2 模型在(6.08°N, 94.6°E)反演得到 13 个潮汐纬向分量

Tab. 1 The 13 tidal zonal components at (6.08°N, 94.6°E) obtained via TPXO7.2 model inversion

潮汐种类	潮汐分量	潮汐速度振幅/($10^{-2} \text{ m}\cdot\text{s}^{-1}$)
半日潮	M_2	5.991 9
	S_2	3.176 7
	N_2	1.516 6
	K_2	0.803 6
全日潮	K_1	1.061 9
	O_1	0.325 4
	P_1	0.369 4
其他混合潮	Q_1	0.017 2
	M_f	0.005 1
	M_m	0.000 5
	M_4	0.016 0
	M_{S4}	-0.000 1
	M_{N4}	0.000 6

表 2 数据库的模拟数据

Tab. 2 Simulation data of the database

序号	潮汐速度振幅/($10^{-2} \text{ m}\cdot\text{s}^{-1}$)	东经/(°)	北纬/(°)	水深/m	时间/h	振幅/m
1	8.12	95.67	6.08	-1 335	6.22	32.01
2	9.45	96.65	6.08	-1 333	34.19	72.58
3	10.08	96.25	6.08	-1 343	27.75	78.45

续表

序号	潮汐速度振幅/($10^{-2} \text{ m}\cdot\text{s}^{-1}$)	东经/(°)	北纬/(°)	水深/m	时间/h	振幅/m
.....
7582	10.29	95.77	6.08	-1 303	9.69	62.60
7583	5.70	96.36	6.08	-1 344	39.67	28.37
7584	7.40	96.67	6.08	-1 330	37.04	63.55

2 预测模型

运用支持向量机和大量的模拟数据建立安达曼海南部内孤立波传播预测模型。通过模型来预测内孤立波的传播过程,包括传播时间、位置和振幅。把数据库中的 7 584 个模拟数据分为两部分,75%为训练数据(序号 1~5688),25%为检验数据(序号 5689~7584)。为衡量模型预测的性能,采用相关系数 R (correlation coefficient),平均绝对误差 M_{AE} (mean absolute error)和平均绝对百分比误差 M_{APE} (mean absolute percentage error)作为模型的评价指标。上述评价指标 R 值越大, M_{AE} 和 M_{APE} 值越小,则预测效果越好。其表达式如下:

$$R = \frac{\sum_{i=1}^n [(y_{\text{true}}(i) - \bar{y}_{\text{true}})(y_{\text{pred}}(i) - \bar{y}_{\text{pred}})]}{\sqrt{\sum_{i=1}^n (y_{\text{true}}(i) - \bar{y}_{\text{true}})^2 \sum_{i=1}^n (y_{\text{pred}}(i) - \bar{y}_{\text{pred}})^2}}, \quad (1)$$

$$M_{AE} = \frac{1}{n} \sum_{i=1}^n |y_{\text{true}}(i) - y_{\text{pred}}(i)|, \quad (2)$$

$$M_{APE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_{\text{true}}(i) - y_{\text{pred}}(i)}{y_{\text{true}}(i)} \right| \times 100\%. \quad (3)$$

式中: n 为预测个数, y_{true} 为真实值, y_{pred} 为模型预测值。

支持向量机(support vector machine, 简称 SVM)由 Cortes 和 Vapnik 于 1995 年正式提出^[13]。它是一种基于统计学习理论的有监督机器学习方法,被广泛地运用在模式识别、函数拟合和时间序列估计等

领域,进行数据分类^[14]、回归和预测^[15-16]研究。本文采用的 SVM 参数有: SVM 设置类型(-s),核函数(-t),惩罚系数(-c),核函数中的 gamma 函数(-g), epsilon-SVR 中的损失函数(-p)。

2.1 内孤立波时间预测模型

把内孤立波到达位置的经度、水深和潮速振幅 3 个参量作为模型输入参量,内孤立波到达时间作为模型输出。内孤立波时间预测模型确定 SVM 设置类型-s 4, nu-SVR 是支持向量回归机,由于 epsilon-SVR 需要事先确定参数,然而在某些情况下选择合适的参数却不是一件容易的事情,而 nu-SVR 能够自动计算参数。经过多次试验和对比验证,确定核函数设置类型-t 0,线性核函数主要用于线性可分的情形。参数少,速度快,对于一般数据,预测效果比较理想。

对时间预测模型检验结果绘制成时间数据散点图和偏差概率图,并计算散点图的相关系数、平均绝对百分比误差及平均绝对误差,如图 2 所示。通过两者比较可以得知基于 SVM 的内孤立波时间预测模型的 M_{APE} 为 1.17%, M_{AE} 为 0.22 h。从图 2 中可以看出数据点在散点图中分布比较集中, R 为 0.999 5,相关度较高。时间数据偏差在 0 h 附近的数据出现概率最高,主要集中在-0.4 h 到 0.4 h 之间,出现概率为 0.893,预测效果较理想。

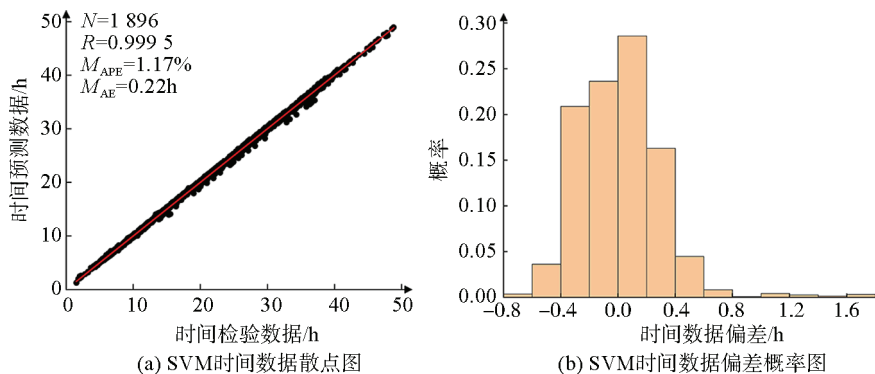


图 2 SVM 时间数据散点图和偏差概率图

Fig. 2 SVM time data scatter diagram and deviation probability chart

2.2 内孤立波位置预测模型

把时间、水深和潮速振幅 3 个参量作为模型输入，到达位置的经度作为模型输出。内孤立波位置预测模型和时间预测模型的设置是一样的，SVM 设置类型-s 4，核函数设置类型-t 0。经过多次试验和对比验证，这个设置不仅适合于时间预测，也同样适合于到达位置预测。模型检验结果的时间数据散点图和偏差概率

图如图 3 所示。通过两者对比可以得出基于 SVM 的内孤立波位置预测模型的 M_{APE} 为 0.016%， M_{AE} 为 0.015°。从图 3 中可以看出虽然有个别数据点分散，但是大部分数据点在散点图中分布较为集中，相关系数 R 为 0.999 1，相关度较高。经度数据偏差在-0.01°附近的数据出现概率最高，主要集中在-0.02°到 0.03°之间，出现概率为 0.865，预测结果比较好。

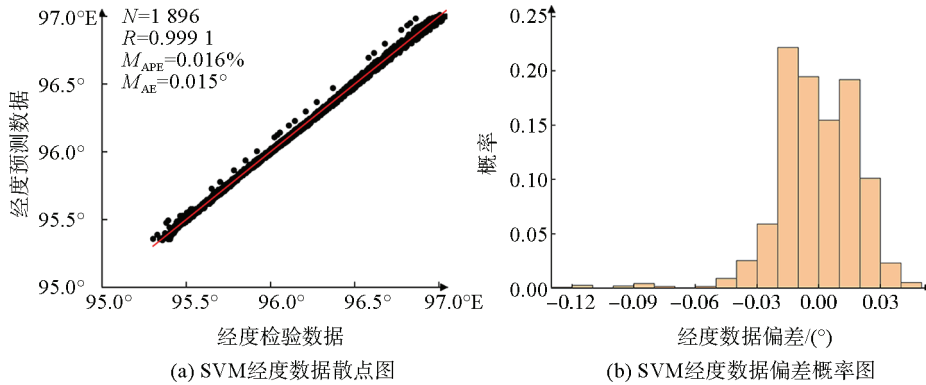


图 3 SVM 经度数据散点图和偏差概率图

Fig. 3 SVM longitude data scatter diagram and deviation probability chart

2.3 内孤立波振幅预测模型

把时间、经度、水深和潮速振幅 4 个参量作为模型输入，振幅作为模型输出。内孤立波振幅预测模型确定 SVM 设置类型-s 3，epsilon-SVR 对应不敏感损失函数 ϵ 的支持向量回归机，对样本点来说，存在着一个不为目标函数提供任何损失值的区域。经过多次试验和对比验证，确定 epsilon-SVR 中的损失函数 $p=0.05$ ，核函数设置类型-t 2，RBF 核函数因为其追踪性能好且无记忆性，通过调节核参数更能反映训练数据的范围分布，预测效果较好。通过多次参数寻优得到惩罚系数 $C_{best}=0.03125$ ，RBF 核函数中的

gamma 函数 $g_{best}=8$ 。

检验模型结果见绘制时间数据散点图和偏差概率图，并计算散点图的相关系数、平均绝对百分比误差及平均绝对误差，如图 4 所示。通过计算得出基于 SVM 的内孤立波振幅预测模型的 M_{APE} 为 10.00%， M_{AE} 为 3.88 m。从图 4 中可以看出数据点在散点图中分布有点分散，但是整体上是较为集中的， $R=0.9758$ ，相关度较高。数据偏差在 0 m 附近的数据出现概率最高，数据偏差比较分散，主要集中在-8 m 到 8 m 之间，出现概率为 0.886，预测结果是可以接受的。

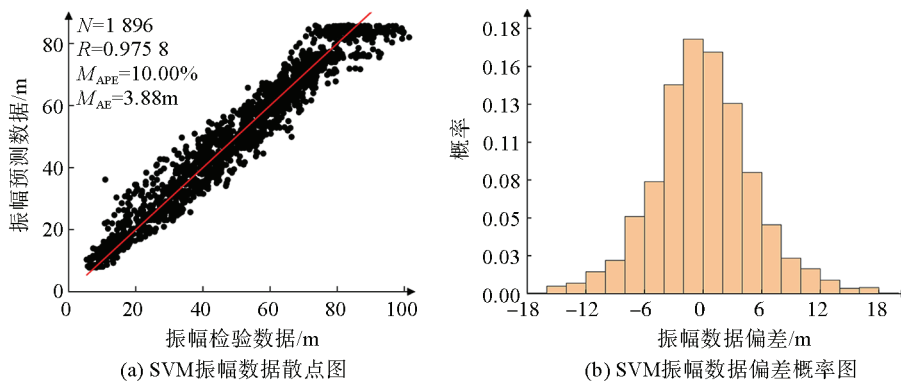


图 4 SVM 振幅预测数据散点图和偏差概率图

Fig. 4 SVM amplitude prediction data scatter diagram and deviation probability chart

3 模型经度验证

在安达海南部海域(5°N~7°N, 94°E~100°E), 收集了 2018 全年的 Sentinel-1A SAR 图像。在所有的图像中总共筛选出 3 景匹配的图像, 如图 5 所示。

从 3 景图像中选择 5 条明显的内孤立波, 确定其

位置后, 通过 ETOPO1 获得红点位置的水深, 通过 TPXO7.2 获得内孤立波生成源处的潮速振幅, 如表 3 所示。运用 SVM 把 7 854 个模拟数据作为训练集进行模型训练, 然后运用训练好的模型来预测图像中的传播时间、位置和振幅, 最终把预测数据与 SAR 图像数据做对比。

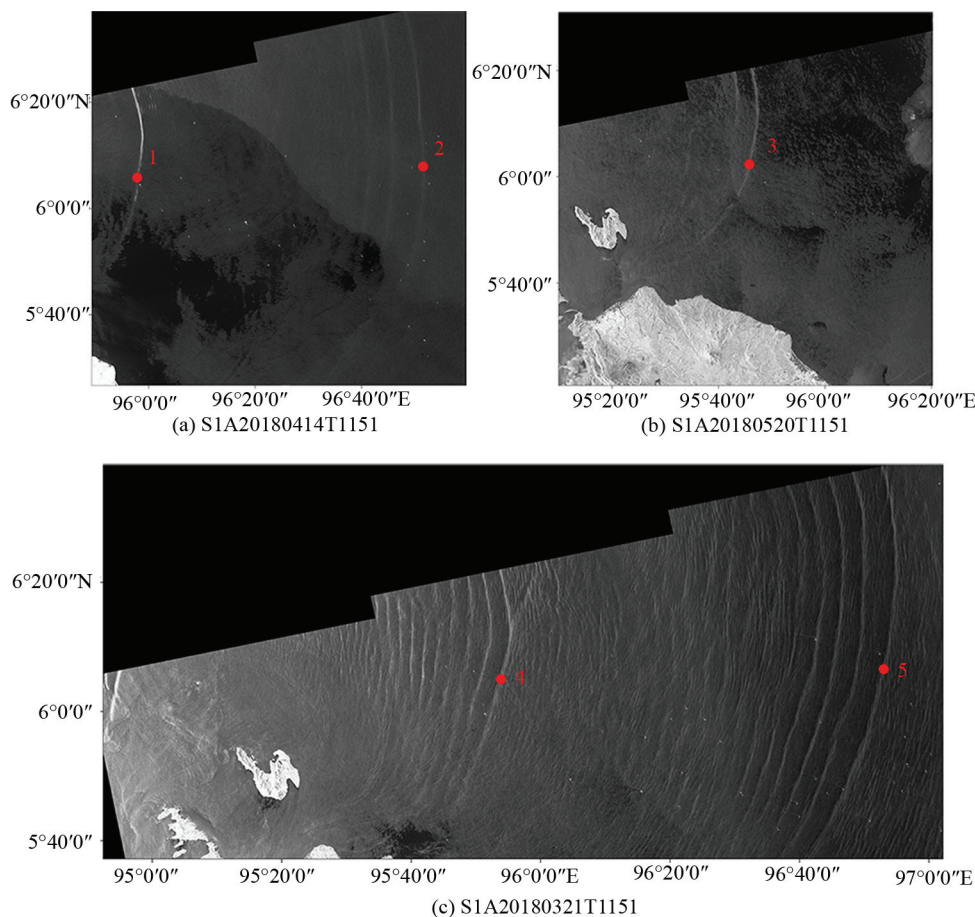


图 5 安达曼海南部 Sentinel-1A SAR 图像

Fig. 5 Sentinel-1A SAR images of the southern Andaman Sea

注: 图中数字与表 3 中数字相对应, 图中红点为表 3 数据的提取位置。

表 3 SAR 图像数据
Tab. 3 SAR image data

序号	图像编号	北纬/(°)	东经/(°)	水深/m	生成源潮速振幅/(10 ⁻² m·s ⁻¹)
1	S1A_IW_GRDH_1SDV_20180414T1151	6.081	95.96	1 343	7.55
2	S1A_IW_GRDH_1SDV_20180414T1151	6.081	96.86	1 274	6.68
3	S1A_IW_GRDH_1SDV_20180520T1151	6.081	95.77	1 330	7.22
4	S1A_IW_GRDH_1SDV_20180321T1151	6.080	95.90	1 346	9.82
5	S1A_IW_GRDH_1SDV_20180321T1151	6.080	96.87	1 274	10.05

3.1 内孤立波到达时间预测

把经度、水深和潮速振幅 3 个参量作为模型输

入, 时间作为模型输出。通过 SVM 的内孤立波时间预测模型, 得到时间预测数据, 如表 4 所示。图 6 是

SVM 预测的时间数据和图像时间数据对比图。圆形代表通过 SVM 时间预测模型预测的时间数据, 正方形代表 SAR 图像的时间数据。通过比较可得, SVM 模型预测的时间数据和 SAR 图像时间数据的 M_{APE} 为 8.43%, M_{AE} 为 1.00 h。

表 4 SVM 时间预测结果

Tab. 4 SVM time prediction results

序号	SVM 预测时间数据 (世界时)	SAR 图像时间数据 (世界时)
1	10.49	11.85
2	11.28	11.85
3	12.02	11.85
4	12.48	11.85
5	14.11	11.85

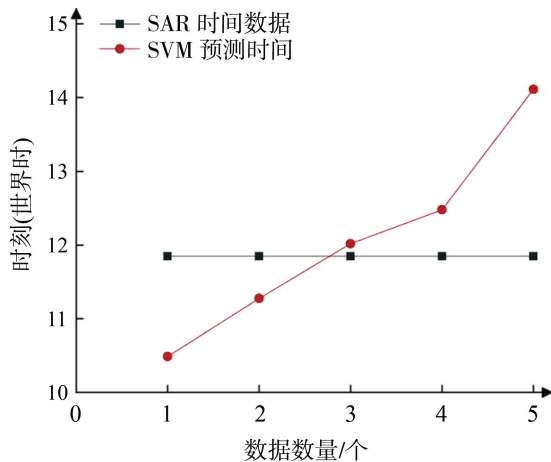


图 6 SVM 预测时间数据和 SAR 图像时间数据对比图
Fig. 6 Comparison of SVM prediction time data and SAR image time data

3.2 内孤立波位置预测模型结果

本文把时间、水深和潮速振幅 3 个参量作为模型输入, 经度作为模型输出。表 5 显示了基于 SVM 的内孤立波位置预测模型的经度预测数据和 SAR 图像经度数据。经度预测数据和图像经度数据对比图如图 7 所示。从图 7 可以得出, 圆形代表通过 SVM 模型得到的预测经度数据, 正方形代表 SAR 图像经度数据。通过对比得知, SVM 模型预测经度和 SAR 图像经度的 M_{APE} 为 0.072%, M_{AE} 为 0.069°。图 7 中的两条曲线拟合程度较好, 预测结果较理想。

表 5 SVM 经度预测结果

Tab. 5 SVM longitude prediction results

序号	SVM 预测内孤立波位置 经度(东经)	SAR 图像内孤立波位置 经度(东经)
1	96.05°	95.96°
2	96.90°	96.86°
3	95.76°	95.77°
4	95.85°	95.90°
5	96.71°	96.87°

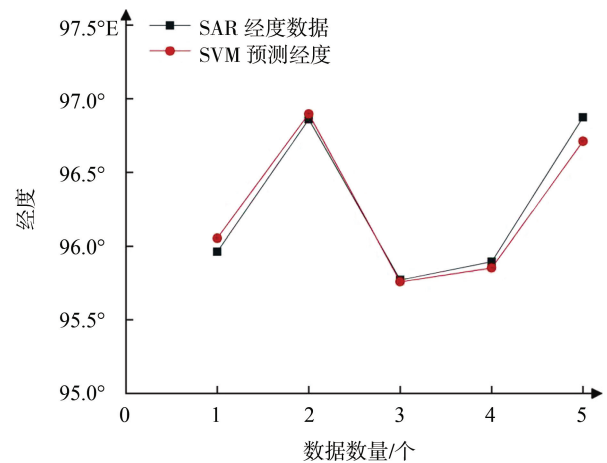


图 7 SVM 预测经度数据和图像经度数据对比图
Fig. 7 Comparison of SVM prediction longitude data and SAR image longitude data

3.3 内孤立波振幅预测模型结果

把时间、经度、水深和潮速振幅 4 个参量作为模型输入, 振幅作为模型输出。通过 SVM 的内孤立波振幅预测模型, 得到振幅预测数据, 如表 5 所示。SVM 振幅预测数据的最大振幅为 84.98 m, 最小振幅为 23.80 m, 平均振幅为 56.26 m。从表 6 中可以看出, SVM 振幅预测数据的最大值在农历二月初五。这一天的潮汐速度最大, 并且内孤立波经度也最大。SVM 振幅预测数据的最小值在农历四月初六。这一天的潮汐速度较小, 而且内孤立波经度最小。由于没有与遥感数据相匹配的现场实测数据, 所以将 SVM 振幅预测的结果与历史的实测数据进行对比。Osborne 等^[4]在安达曼海南部 1 093 m 深的水域, (6°53'N, 97°04'E)位置处测量到振幅为 60 m 的内孤立波。实测位置和预测位置接近, 预测的平均振幅和实测振幅差距较小, 所以内孤立波振幅预测结果是合理的。

表 6 SVM 振幅预测结果

Tab. 6 SVM amplitude prediction results

序号	日期	SVM 振幅预测数据/m	东经/(°)	生成源潮速振幅/(10 ⁻² m·s ⁻¹)
1	农历二月二十九	48.36	95.96	7.55
2	农历二月二十九	59.40	96.86	6.68
3	农历四月初六	23.80	95.77	7.22
4	农历二月初五	64.96	95.90	9.82
5	农历二月初五	84.98	96.87	10.05

4 结论

本文提出了一种基于机器学习来预测内孤立波传播的新思路,并建立了基于 SVM 的安达曼海南部内孤立波传播预测模型。首先运用全球潮汐模式 TPXO7.2 得到潮汐速度,以潮汐速度为初始强迫场构造出不同潮汐速度下的 MITgcm 模式的内孤立波数值计算模型。通过 MITgcm 模拟计算了潮汐速度 0.05~0.1 m/s 的内孤立波生成与传播过程,总共提取了 7 584 条内孤立波。然后把 7 584 个模拟数据分为两部分,5 688 个数据为训练数据,1 896 个数据为检验数据。采用 SVM 方法建立了安达曼海南部内孤立波传播预测模型。在模型建立的过程中,基于 SVM 的内孤立波传播预测模型的时间模型的 R 为 0.999 5, M_{APE} 为 1.17%, M_{AE} 为 0.22 h。位置模型的 R 为 0.999 1, M_{APE} 为 0.016%, M_{AE} 为 0.015°。振幅模型的 R 为 0.975 8, M_{APE} 为 10.00%, M_{AE} 为 3.88 m。

然后运用 Sentinel-1A 的 SAR 图像中的数据来验证模型的准确性。通过与 SAR 图像对比可得, SVM 模型预测的时间数据和图像时间数据的 M_{APE} 为 8.43%, M_{AE} 为 1.00 h。SVM 模型预测的经度和图像经度的 M_{APE} 为 0.072%, M_{AE} 为 0.069°。SVM 的振幅预测数据平均振幅为 56.26 m,历史上的实测数据振幅为 60 m,两者比较接近。因此, SVM 的安达曼海南部内孤立波传播预测模型是有效的,可为石油勘探和海洋决策部门预警提供参考。

参考文献:

[1] 方欣华,杜涛. 海洋内波基础和中国海内波[M]. 青岛: 中国海洋大学出版社, 2005: 1-4.
FANG Xinhua, DU Tao. Fundamentals of oceanic internal waves and internal waves in the China seas[M]. Qingdao: China Ocean University Press, 2005: 1-4.

[2] 段亚男. 基于 ARM 嵌入式技术监测海洋内孤立波传播特征的系统设计[D]. 青岛: 青岛理工大学, 2013.
DUAN Ya'nan. Designing based on ARM embedded technology to detect propagation characteristics of in-

ternal solitary waves[D]. Qingdao: Qingdao University of Technology, 2013.

[3] 蔺飞龙. 南海北部内波特征研究及机理分析[D]. 北京: 中国科学院大学, 2014.
LIN Feilong. Characteristics and mechanism analysis of internal waves in the northern South China Sea[D]. Beijing: University of Chinese Academy of Sciences, 2014.

[4] OSBORNE A R, BURCH T L. Internal solitons in the Andaman sea[J]. Science, 1980, 208(4443): 451-460.

[5] HYDER P, JEANS D R G, CAUQUIL E, et al. Observations and predictability of internal solitons in the northern Andaman Sea[J]. Applied Ocean Research, 2005, 27(1): 1-11.

[6] ALPERS W, HENG Wangchen, LIM Hock. Observation of internal waves in the Andaman Sea by ERS SAR[J]. European Space Agency, (Special Publication) ESA SP, 1997, 414(3): 1287-1291.

[7] JACKSON C. Internal wave detection using the moderate resolution imaging spectroradiometer (MODIS)[J]. Journal of Geophysical Research: Oceans, 2007, 112: C11012.

[8] DA SILVA J C B, MAGALHAES J M. Internal solitons in the Andaman Sea: a new look at an old problem[C]// Proceedings of the Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions. Edinburgh, UK: SPIE Remote Sensing, 2016, 9999(999907).

[9] SHIMIZU K, NAKAYAMA K. Effects of topography and earth's rotation on the oblique interaction of internal solitary-like waves in the Andaman Sea[J]. Journal of Geophysical Research: Oceans, 2017, 122(9): 7449-7465.

[10] SUN Lina, ZHANG Jie, MENG Junmin. A study of the spatial-temporal distribution and propagation characteristics of internal waves in the Andaman Sea using MODIS[J]. Acta Oceanologica Sinica, 2019, 38(7): 121-128.

[11] RAJU N J, DASH M K, DEY S P, et al. Potential generation sites of internal solitary waves and their propagation characteristics in the Andaman Sea-a study based on MODIS true-colour and SAR observations.[J]. Environmental monitoring and assessment, 2019, 191(Suppl 3): 809.

- [12] ZHANG Z, FRINGER O B, RAMP S R. Three-dimensional, nonhydrostatic numerical simulation of nonlinear internal wave generation and propagation in the South China Sea[J]. *Journal of Geophysical Research Oceans*, 2011, 116(C5): C05022.
- [13] CORTES C, VAPNIK V N. Support-Vector networks[J]. *Machine Learning*, 1995, 20(3): 273-297.
- [14] 张锦水, 何春阳, 潘耀忠, 等. 基于 SVM 的多源信息复合的高空间分辨率遥感数据分类研究[J]. *遥感学报*, 2006, 10(1): 49-57.
ZHANG Jinshui, HE Chunyang, PAN Yaozhong, et al. The high spatial resolution RS image classification based on SVM method with the multi-source data[J]. *Journal of Remote Sensing*, 2006, 10(1): 49-57.
- [15] 吴恒, 余秋鹏. 基于 SVM 预测电动汽车并网局部电压稳定的研究[J]. *电工电气*, 2019(10): 9-13.
WU Heng, YU Qiupeng. Research on prediction of local voltage stability of electric vehicles based on SVM[J]. *Electrotechnics Electric*, 2019(10): 9-13.
- [16] 徐娟, 卞良. 基于 SVM 的中文垃圾邮件预测系统研究[J]. *数字技术与应用*, 2020, 38(355): 48-49.
XU Juan, BIAN Liang. Research on Chinese spam prediction system based on SVM[J]. *Digital Technology and Application*, 2020, 38(355): 48-49.

The prediction of internal solitary wave propagation characteristics in the southern Andaman Sea based on a support vector machine

LU Ke-xiao, WANG Jing, WEI Xin

(College of Physics and Optoelectronic Engineering, Ocean University of China, Qingdao 266100, China)

Received: Nov. 5, 2020

Key words: internal solitary waves; MITgcm numerical simulation; support vector machine(SVM); prediction

Abstract: Internal solitary wave (ISW) is a special type of internal wave. It occurs in density-stable stratified seawater and have strong randomness. Their occurrence location, arrival time, amplitude, and other parameters are greatly affected by hydrology and other aspects of the external environment. Therefore, accurate prediction of internal solitary waves propagation is a recognized challenge. In this paper, a method is proposed to use the Massachusetts Institute of Technology general circulation model (MITgcm) of ISWs to calculate abundant simulation data and establish a database. A prediction model of ISWs propagation in the southern Andaman Sea based on a support vector machine (SVM) is established using machine learning. Finally, the Sentinel-1 synthetic aperture radar (SAR) images in the southern Andaman Sea are used to test the results of the ISWs propagation prediction model. The results demonstrate that the mean absolute percentage error of time (location) predicted by the prediction model of ISWs propagation time (arrival location) based on an SVM is 8.43% (0.072%), and the mean absolute error is 1.00 h (0.069°). The predicted amplitude range of the ISWs amplitude prediction model based on an SVM is 23.80~84.98 m.

(本文编辑: 丛培秀)