

[海浪智能快速预报模型演示软件]

说明书

[V1.0]

申报者姓名： 王志亮 姚宇 李江夏

软件首次发表日期：2024 年 10 月 12 日

目录

1 引言	1
1.1 开发目的	1
1.2 开发背景	1
1.3 相关定义	1
2 软件概述	2
2.1 逻辑与功能	2
2.2 开发运行环境	3
2.3 编译环境	4
2.4 启动与运行	5
3 案例展示—以中国南海波高预测为例	7
3.1 数据准备	7
3.2 运行演示软件	9
3.3 绘制年平均有效波高图	10
3.4 调取所需年份有效波高图	12
附件 1：预测模型验证报告	14
附件 2：训练模型 Python 代码	19
附件 3：预测模型 Python 代码	31

1 引言

1.1 开发目的

本软件的开发目的是将人工智能技术的方法和思路应用到海浪的模拟与预报中，建立数据驱动的海浪智能预报模型，并提供一种波浪场可视化方法，系统性实现海洋波浪场的智能快速预报。以便于用户可随时预测并获取目标区域未来的波高或周期分布情况以及指定经纬度地区附近的年均有效波高或周期值，类似于天气预报软件，可为渔民、出海搜救人员、科研工作者等有特定需求群体提供一定的数据参考价值。

1.2 开发背景

海洋表面重力波（以下简称波浪）具有很强的非线性，对海洋工程活动、海上作业和运输都有重大影响，实现对波浪进行精确模拟和预报具有重要的现实意义。而目前常用的数值预报模型是物理规律驱动数值逼近模型，通过迭代计算求解物理方程实现波浪场预报，这种传统的数值模式预报难以对波浪的非线性复杂过程进行准确模拟，导致预报精度受限。随着人工智能技术的不断发展，其展现的强大非线性学习和拟合能力可以有效地利用海洋大数据中固有的时空属性，以提高对波浪变化预报的准确性和时效性。因此，本软件基于大数据驱动的智能预报模型，利用深度机器学习方法从大量历史风浪数据中学习海浪的时空演化规律，从而实现波浪场预报。

1.3 相关定义

波浪场：描述波浪要素的空间分布特征；

有效波高 (H_s)：指在任一个由 n 个波浪组成的波群中，将波列中的波高由大到小依次排列，确定前 $n/3$ 个波为有效波，有效波的波高则等于这 $n/3$ 个波的平均波高；

年均有效波高：某一年有效波高的算术平均值；

MLP 模型：多层感知器（MLP, Multilayer Perceptron）是一种前馈人工神经网络模型，其将输入的多个数据集映射到单一的输出的数据集上。其构成由一个输入层、多个隐藏层和一个输出层组成，最简单的 MLP 模型只含一个隐层，即三层的结构。层与层之间是全连接的（上一层的任何一个神经元与下一层的所有神经元都有连接），最底层是输入层，中间是隐藏层，最后是输出层。

2 软件概述

2.1 逻辑与功能

海浪智能快速预报模型是利用 Python 库中的 Scikit-Learn 0.19.2 版进行开发的，Scikit-Learn 中有多种回归算法，包括线性回归、基于树的回归器（即决策树回归器、随机森林回归器）和 MLP 回归器。波浪预测是一个非线性问题，因此线性回归并不适合这一应用。在基于树的模型中，首先将预测空间划分为若干不同的非重叠区域，对于属于某个区域的每个观测值，都会做出相应的预测，预测值就是该观测值的响应值的平均值，也不适用于波浪的时空预测。因此，当前的应用选择了 MLP 回归器。

预报模型的逻辑架构如图 1 所示，首先收集获取驱动模型所需的大数据，该数据可通过现场实测、遥感卫星图像分析、再分析数据库

等多种方式获取，对于遥感卫星图像分析，首先要获取含有波浪特征的遥感图像数据，对其进行预处理包括大气校正和辐射校正，后再经几何校正和图像处理，提取波浪的时间信息和空间信息构建数据样本库，进而得到模型的训练数据集，由于演示案例中现有开放获取的南海遥感数据相对有限且敏感，所以本软件暂时基于 FVCOM-SWAVE 数值模拟输出数据对模型进行验证，目的在于测试模型的可靠性和预测的可行性，重点在于提供一种预报的方法和思路。随后对输入有效波高数据进行 MLP 神经网络模型训练、验证和测试，当训练验证误差最小时，即可采用该最优参数的模型进行预测，最后将预测结果通过演示软件进行可视化展示，以便于用户便捷查看数据。

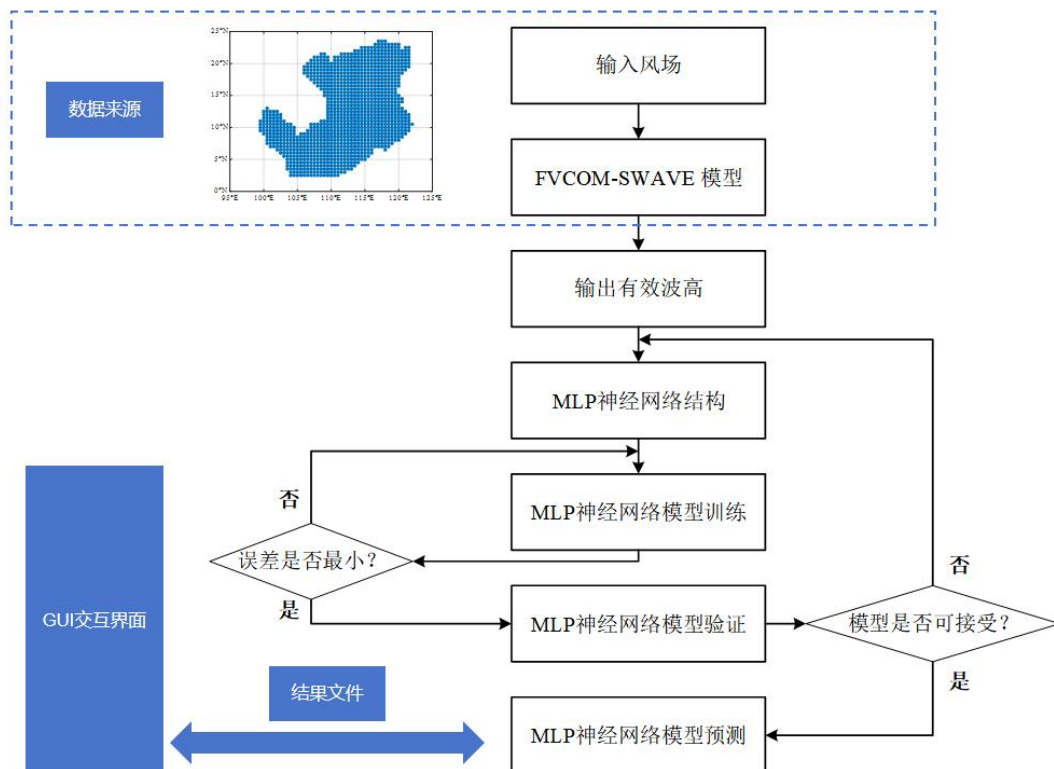


图 1 海浪波高智能快速预报模型逻辑架构

2.2 开发运行环境

软件开发环境：windows10 操作系统；MATLAB R2022a, Python 3.9（64-bit），PyCharm 2021 编译软件。

硬件运行环境：11th Gen Intel(R) Core(TM) i5-11400 @ 2.60GHz 2.59 GHz CPU, 16GB 机带 RAM, 1TB 固态硬盘, 1Mbps 网络带宽；

2.3 编译环境

预测模型基于 Python 3.9（64-bit）使用 PyCharm 2021 编译软件进行开发编译，演示软件使用 MATLAB R2022a 进行开发编译，需要使用相同或更高版本软件进行开发编译。

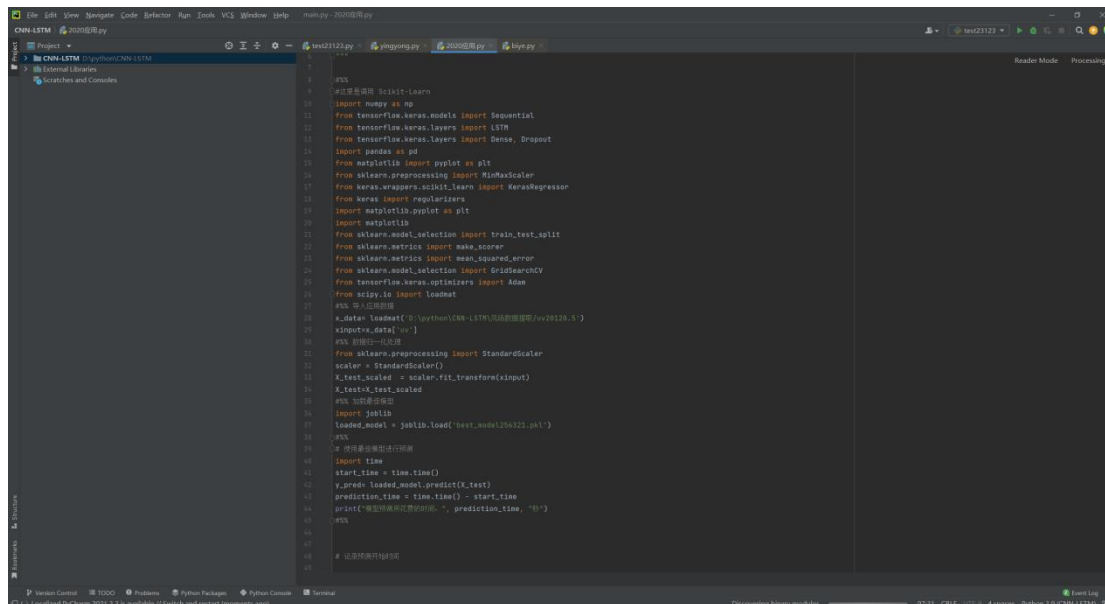


图 2 预测模型 Python 代码编译界面

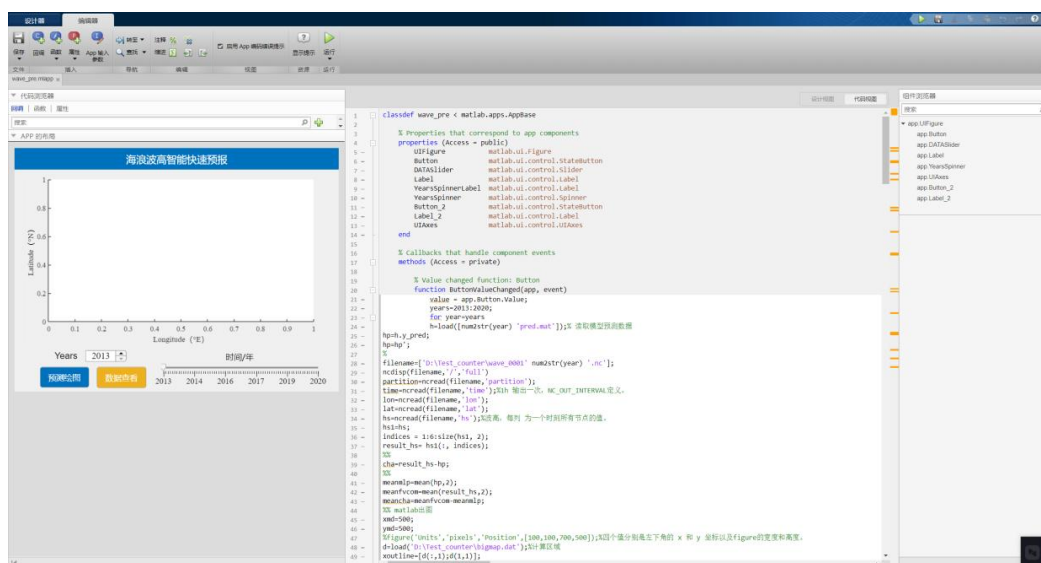


图 3 演示软件 matlab 代码编译及设计界面窗口

2.4 启动与运行

训练模型的代码基于 PyCharm 2021 平台运行，点击 test.py 文件后，将启动 PyCharm 2021 软件，同时进入如下软件主界面，点击运行按钮即可进行模型训练：

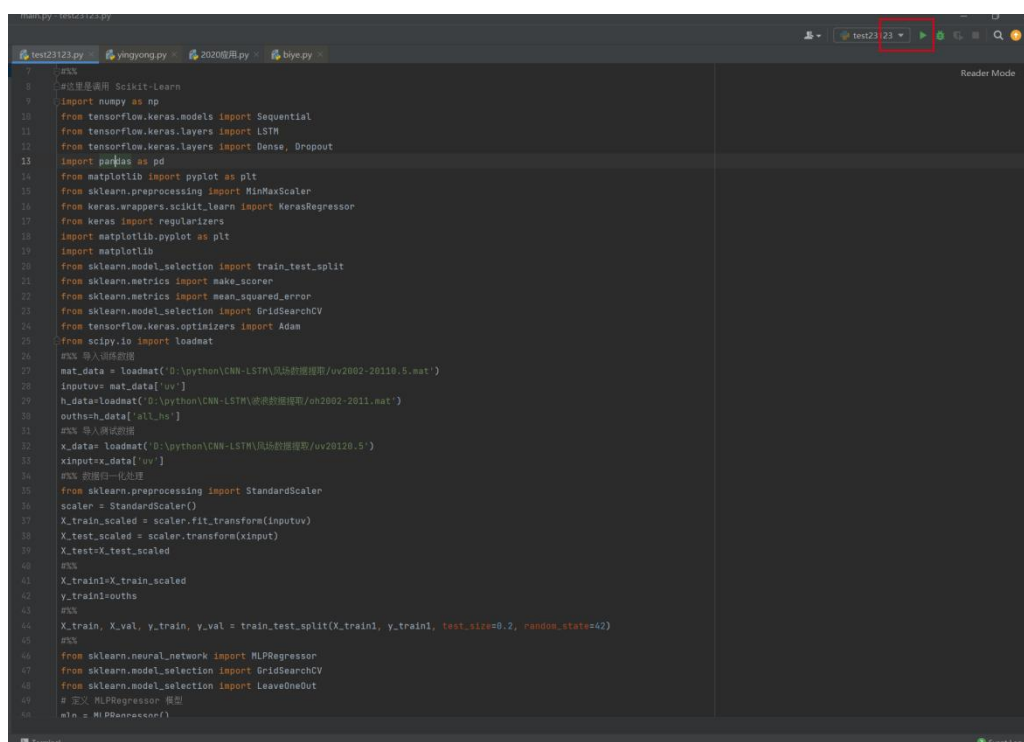


图 4 训练模型代码运行界面

模型训练运行结束后，预测模型的最佳参数文件将以.pkl 格式保存在当前运行文件夹内，如图所示：

2012应用.py	2024/5/28 22:10	PY 文件
best_model.pkl	2024/8/2 21:42	PKL 文件
main.py	2024/5/28 10:31	PY 文件
test.py	2024/6/4 8:42	PY 文件

图 5 最佳训练模型文件保存

预测模型的代码也是基于 PyCharm 2021 平台运行，点击 2012 应用.py 文件后，将启动 PyCharm 2021 软件，同时进入如下软件主界面，点击运行按钮即可进行模型预测：

```

7
8
9 #这里是调用 Scikit-Learn
10 import...
11 #%% 导入应用数据
12 x_data= loadmat('D:\python\CNN-LSTM\浪潮数据源\ov20120.5')
13 input=x_data['ov']
14 #%% 数据归一化处理
15 from sklearn.preprocessing import StandardScaler
16 scaler = StandardScaler()
17 X_test_scaled = scaler.fit_transform(input)
18 X_test-X_test_scaled
19 #%% 加载最佳模型
20 import joblib
21 loaded_model = joblib.load('best_model256321.pkl')
22 #%%
23 # 使用最佳模型进行预测
24
25 import time
26 start_time = time.time()
27 y_pred= loaded_model.predict(X_test)
28 prediction_time = time.time() - start_time
29 print("模型预测所花费的时间: ", prediction_time, "秒")
30 #%%
31
32 # 记录预测开始时间
33
34
35 # 使用最佳模型进行预测
36 y_pred = loaded_model.predict(X_test)
37
38 # 计算预测所花费的时间
39
40
41 #%% 绘制模型应用的误差图
42 import seaborn as sns
43 import matplotlib.pyplot as plt
44 import numpy as np
45 from sklearn.metrics import mean_squared_error, r2_score
46 # 假设这里 y_true 是 y_pred 是您的实例数据
47 h_data=loadmat('D:\python\CNN-LSTM\浪潮数据源\TQh2012.mat')
48 h2012=h_data['h']
49 #%%

```

图 6 预测模型代码运行界面

演示软件基于 MATLAB 平台运行，点击.mlapp 文件后，将启动 MATLAB 软件，同时进入如下软件主界面：

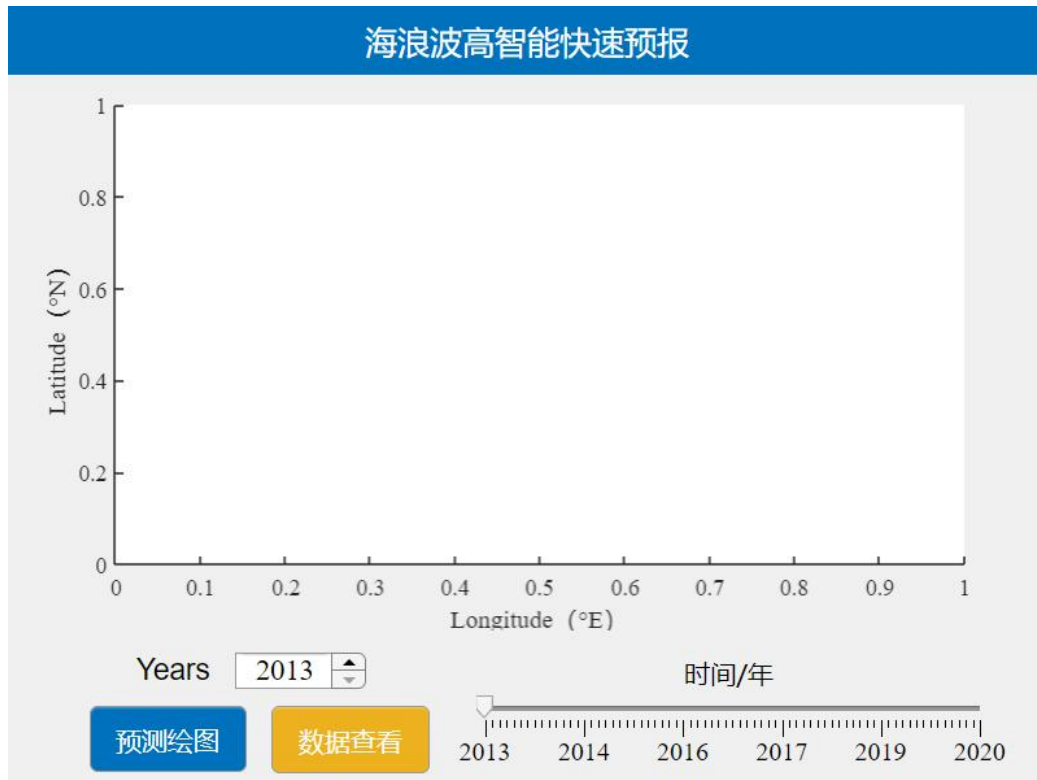


图 7 演示软件启动界面

3 案例展示—以中国南海波高预测为例

3.1 数据准备

本案例以南海 2002-2011 年的波高数据作为模型训练集，2012 年的为模型验证集，经过遥感图像数据分析处理后的波高和风场训练数据以.mat 格式保存在运行文件夹中，如图 8、图 9 所示，基于 MLP 模型进行训练并验证，模型训练过程如图 10 所示，随后应用训练的最佳模型分别对 2013-2020 年的波高进行预测，并输出数据保存为.mat 格式，保存的预测数据如图 11 所示：




















 TQh2002.mat	2024/3/5 11:00	MAT 文件
 TQh2003.mat	2024/3/5 11:00	MAT 文件
 TQh2004.mat	2024/3/5 11:00	MAT 文件
 TQh2005.mat	2024/3/5 11:00	MAT 文件
 TQh2006.mat	2024/3/5 11:00	MAT 文件
 TQh2007.mat	2024/3/5 11:00	MAT 文件
 TQh2008.mat	2024/3/5 11:00	MAT 文件
 TQh2009.mat	2024/3/5 11:00	MAT 文件
 TQh2010.mat	2024/3/5 11:00	MAT 文件
 TQh2011.mat	2024/3/5 11:00	MAT 文件
 TQh2012.mat	2024/3/5 11:05	MAT 文件
 TQh2013.mat	2024/3/22 17:13	MAT 文件
 TQh2014.mat	2024/3/22 17:13	MAT 文件
 TQh2015.mat	2024/3/22 17:13	MAT 文件
 TQh2016.mat	2024/3/22 17:13	MAT 文件
 TQh2017.mat	2024/3/22 17:13	MAT 文件
 TQh2018.mat	2024/3/22 17:13	MAT 文件
 TQh2019.mat	2024/3/22 17:13	MAT 文件
 TQh2020.mat	2024/3/22 17:13	MAT 文件

图 8 2002-2020 年中国南海波浪数据











 uv2002-20110.5.mat	2024/3/5 10:54	MAT 文件
 uv20120.5.mat	2024/3/5 10:57	MAT 文件
 uv20130.5.mat	2024/3/22 17:17	MAT 文件
 uv20140.5.mat	2024/3/22 17:17	MAT 文件
 uv20150.5.mat	2024/3/22 17:17	MAT 文件
 uv20160.5.mat	2024/3/22 17:18	MAT 文件
 uv20170.5.mat	2024/3/22 17:18	MAT 文件
 uv20180.5.mat	2024/3/22 17:18	MAT 文件
 uv20190.5.mat	2024/3/22 17:18	MAT 文件
 uv20200.5.mat	2024/3/22 17:18	MAT 文件

图 9 2002-2020 年中国南海风场数据

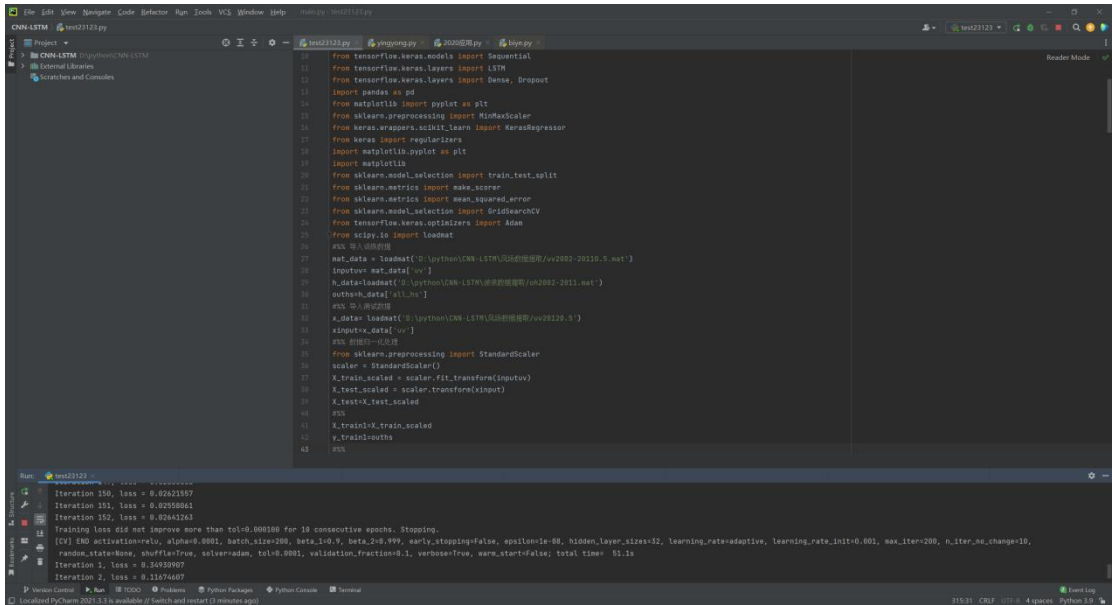


图 10 案例模型预测运行界面

2020pred.mat	2024/3/22 17:27	MAT 文件	16,219 KB
2019pred.mat	2024/3/22 17:27	MAT 文件	16,175 KB
2018pred.mat	2024/3/22 17:26	MAT 文件	16,175 KB
2017pred.mat	2024/3/22 17:26	MAT 文件	16,175 KB
2016pred.mat	2024/3/22 17:26	MAT 文件	16,219 KB
2015pred.mat	2024/3/22 17:25	MAT 文件	16,175 KB
2014pred.mat	2024/3/22 17:25	MAT 文件	16,175 KB
2013pred.mat	2024/3/22 17:25	MAT 文件	16,175 KB

图 11 模型预测后保存为.mat 格式的数据

3.2 运行演示软件

演示软件基于 MATLAB 平台运行，将.mlapp 文件与上述保存的数据放在同一个文件夹中，打开.mlapp 文件点击“运行”命令，进入软件主界面，如图 12 所示：

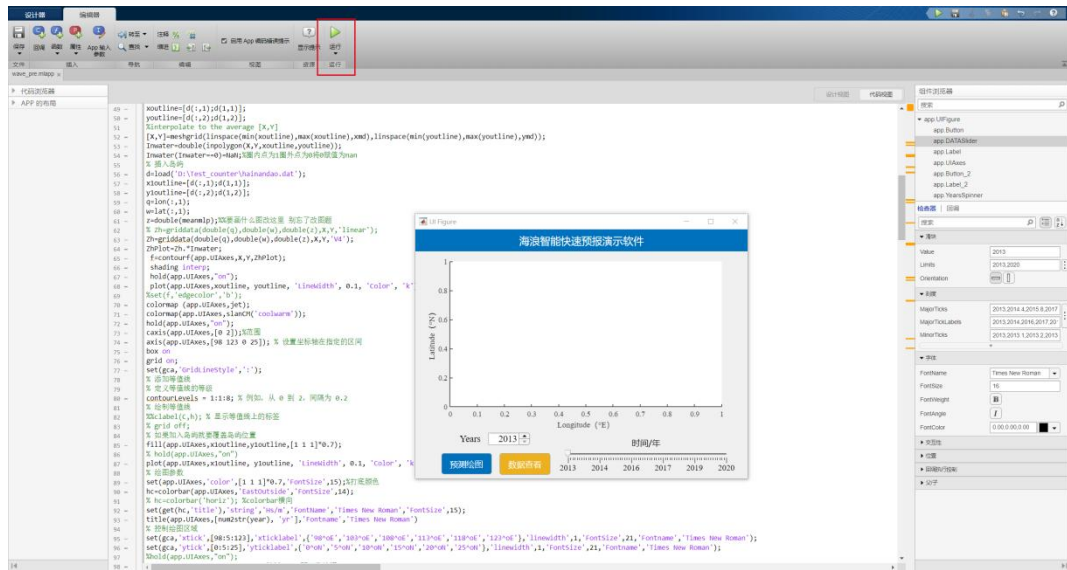


图 12 演示软件成功运行界面

3.3 绘制年平均有效波高图

点击主界面中“预测绘图”按钮，如图 13 所示，可绘制 2013-2020 年的年平均有效波高图，软件成功运行后，如图 14 所示，窗口会依次显示 2013 至 2020 年年均有效波高的云图，运行完成时最终图形窗口显示 2020 年年均有效波高的云图（即预测目标最后时间段的云图），如图 15 所示：

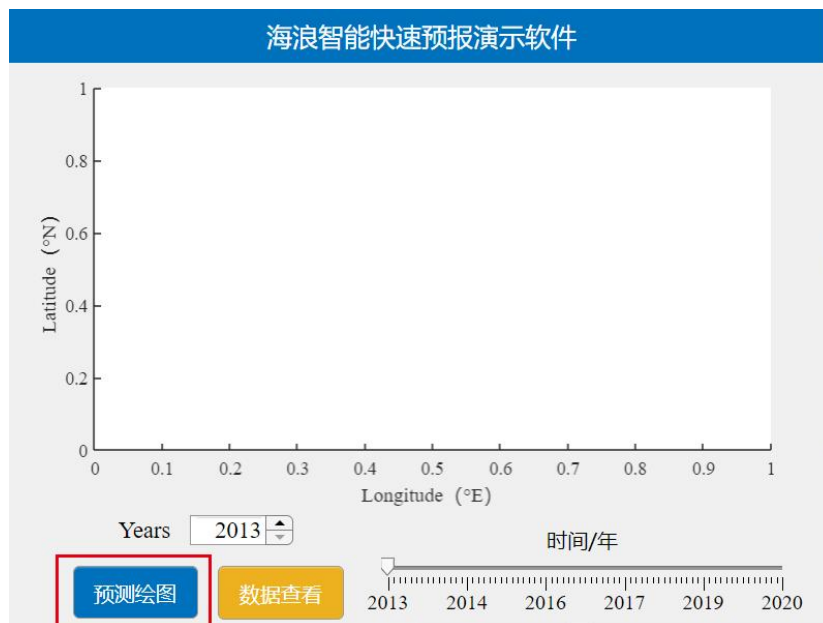


图 13 点击“预测绘图”按钮界面



图 14 程序成功运行时后台命令行窗口显示

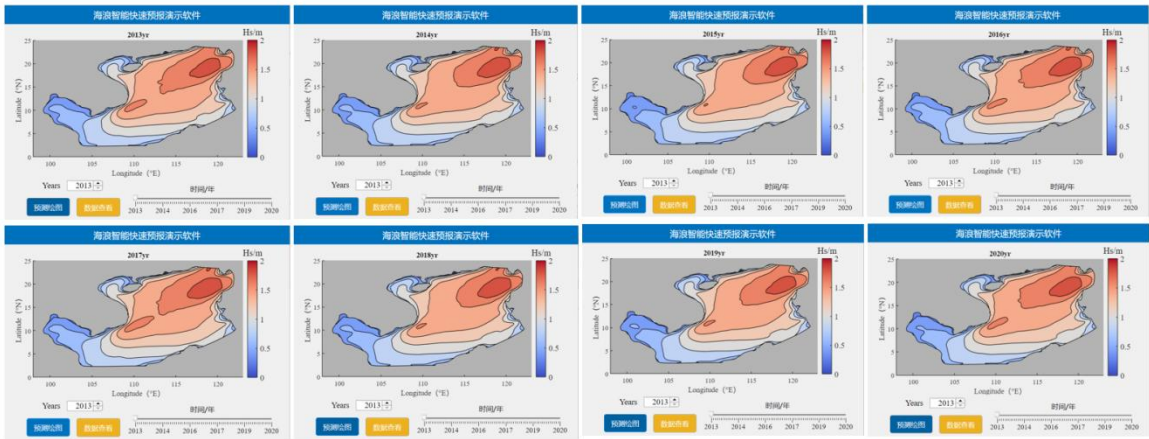


图 15 预测绘图图形窗口云图显示

与此同时，此命令运行完成后会自动按各年份命名，并以 jpg 格式保存在当前程序运行文件夹内，如图 16 所示：

	2020.jpg	2024/7/30 22:51	JPG 图片文件	84 KB
	2019.jpg	2024/7/30 22:50	JPG 图片文件	83 KB
	2018.jpg	2024/7/30 22:49	JPG 图片文件	84 KB
	2017.jpg	2024/7/30 22:49	JPG 图片文件	84 KB
	2016.jpg	2024/7/30 22:48	JPG 图片文件	84 KB
	2015.jpg	2024/7/30 22:48	JPG 图片文件	83 KB
	2014.jpg	2024/7/30 22:47	JPG 图片文件	84 KB
	2013.jpg	2024/7/30 22:47	JPG 图片文件	84 KB

图 16 保存在文件夹中的 jpg 格式图片

另外，点击云图界面内任意一个经纬度位置即可显示该位置的具

体年均有效波高值，例如，经纬度为（117.6°E，15.51°N）在 2020 年的年均有效波高值为 1.583，如图 17 所示：

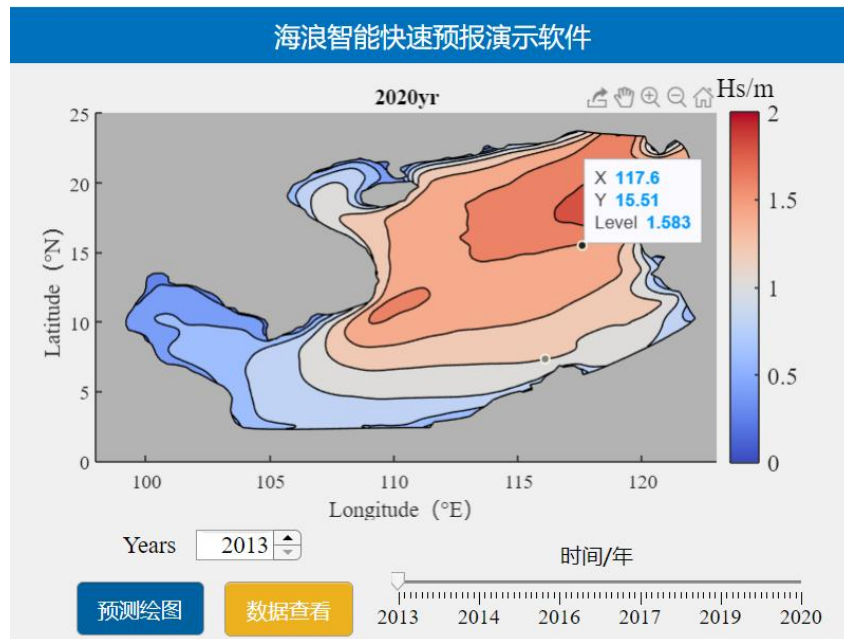


图 17 点击指定经纬度的波高数据显示

3.4 调取所需年份有效波高图

通过“年份选择”按钮选定所需年份数据云图，再点击主界面中“数据查看”按钮，即可查看预测绘图后 2013-2020 年间任意一年的年平均有效波高图，便于满足用户数据需求，如图 18、19 所示。

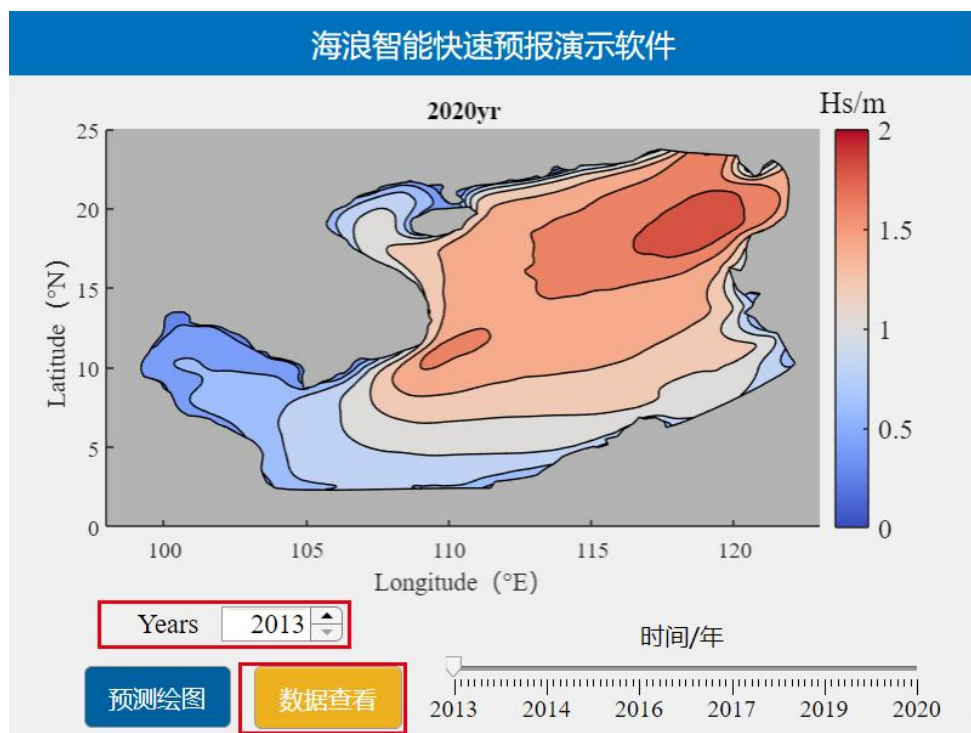


图 18 选定年份后点击“数据查看”按钮界面

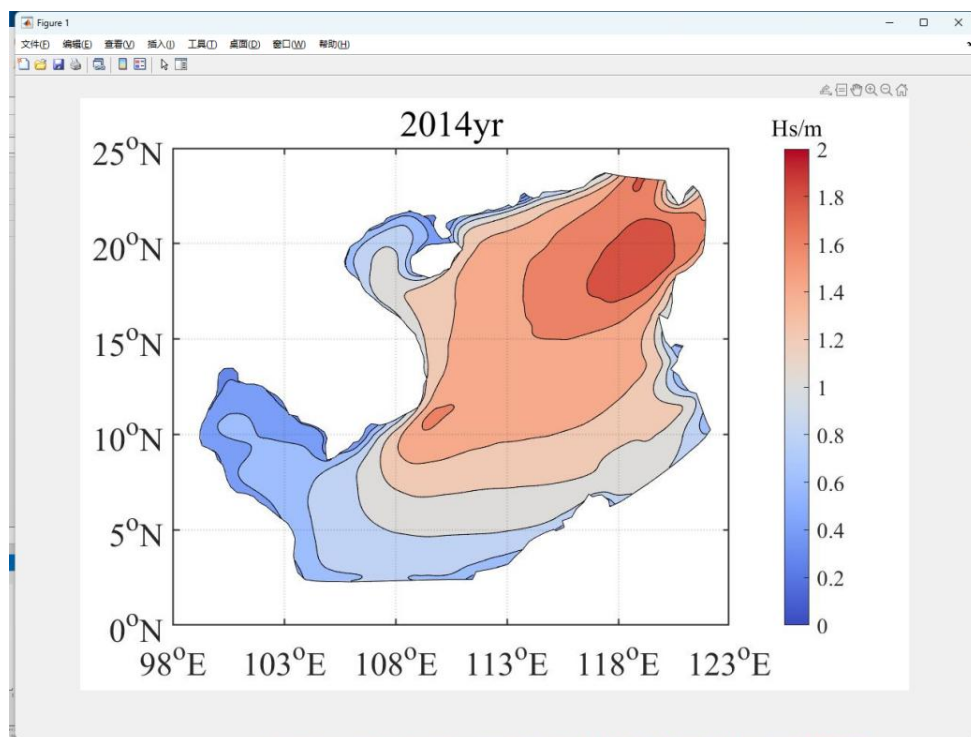


图 19 数据查看操作后弹窗界面显示（以 2014 年为例）

附件 1：预测模型验证报告

模型采用的回归模型评价指标包括决定系数（ R^2 ，coefficient of determination）、均方误差（MSE，Mean Squared Error）以及平均绝对误差（MAE，Mean Absolute Error）。其中，均方误差（MSE）表示模型预测值与实际观测值之间的差异的平方的平均值，该值越大说明模型误差越大。平均绝对误差（MAE）是模型预测值与实际观测值之间的绝对差异的平均值，它能够直观地反映出模型的预测准确程度，其值越大，说明模型误差越大。决定系数（ R^2 ）表示模型拟合数据的程度，其取值范围在 0 到 1 之间，越接近 1 表示模型对数据的拟合程度越好。它们具体计算公式如下：

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_{pi} - y_{oi})^2 \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{pi} - y_{oi})^2} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_{pi} - y_{oi}| \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_{pi} - y_{oi})^2}{\sum_{i=1}^n (y_{pi} - \bar{y}_o)^2} \quad (4)$$

其中， n 是数据总数， y_{pi} 是输出值/预测值， y_{oi} 是目标值/观测值， \bar{y}_o 是测量值的平均值。

在预测模型中，隐藏层个数和神经元数目是对模型性能产生最大影响的参数，这里采用 GridSearchCV 测试了不同隐藏层数量（一层、

两层或三层）和每层神经元数量（32、64、128 和 256）对训练结果的影响，共进行了 12 种组合的实验。设置最大迭代次数为 200，单次传递给程序用以训练的数据为 200，在这里选择 80%的数据进行训练，剩余 20%数据用于验证。此外，每一层均使用了 ReLU 激活函数，其余参数均使用默认值。实验结果表明，具有两层神经网络，每层 256 个神经元的模型训练结果最佳。表 1.1 展示了不同隐藏层个数和每层神经元数量下的 MAE 和 R^2 指标。从结果可以看出，选择神经元数量为 256 的两层神经网络模型表现最好。图 1.1 展示了在最佳参数设置下，损失函数随着迭代次数的变化情况，可以观察到，当迭代次数达到 100 时，优化算法的容差小于 0.0001，说明模型在给定的最大迭代次数前就已经收敛。

表 1.1 神经网络层数对比试验结果

隐藏层	神经元层数数量	MAE	R^2
1	32	0.1994	0.8119
1	64	0.1923	0.8360
1	128	0.1823	0.8495
1	256	0.1770	0.8481
2	(32, 32)	0.1968	0.8158
2	(64, 64)	0.1836	0.8433
2	(128, 128)	0.1688	0.8414
2	(256, 256)	0.1586	0.8650
3	(32, 32, 32)	0.1955	0.8162
3	(64, 64, 64)	0.1809	0.8454
3	(128, 128, 128)	0.1690	0.8557
3	(256, 256, 256)	0.1593	0.8639

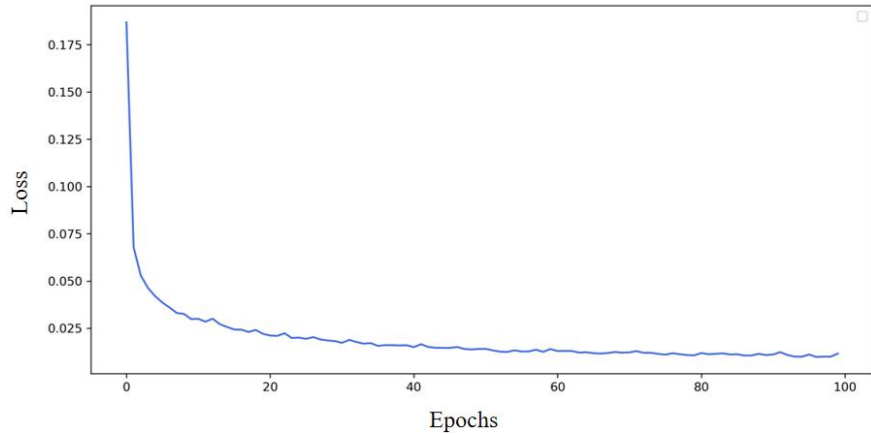


图 1.1 损失系数随代次数增加的变化

模型训练完成后, 利用剩余 20% 的验证数据集进一步评估模型的性能。图 1.2 显示了验证数据集上 (2000-2011) FVCOM-SWAVE 模型和两层 (256, 256) MLP 模型的空间平均有效波高的交叉验证散点图。通过计算得出 R^2 和 RMSE 分别为 0.96 和 0.12 m, 说明所训练的 MLP 神经网络模型在验证数据集上表现出了良好的性能和泛化能力, 能够较为准确地预测中国南海的波高情况。

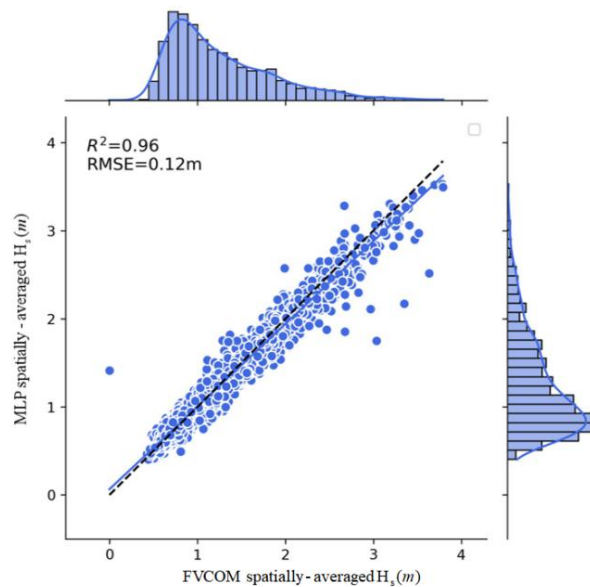


图 1.2 FVCOM-SWAVE 和 MLP 模型验证集上中国南海空间平均有效波高交叉验证散点图 (黑色点划线代表 1:1 线, 蓝实线代表最佳拟合线)

随后, 使用训练好的 MLP 神经网络模型对 2012 年的有效波高进行预测, 并将预测结果与两个实测浮标站和数值模拟的波高结果进

行对比，结果显示 MLP 模型预测的有效波高更加贴合实测波高，如图 1.3 所示。

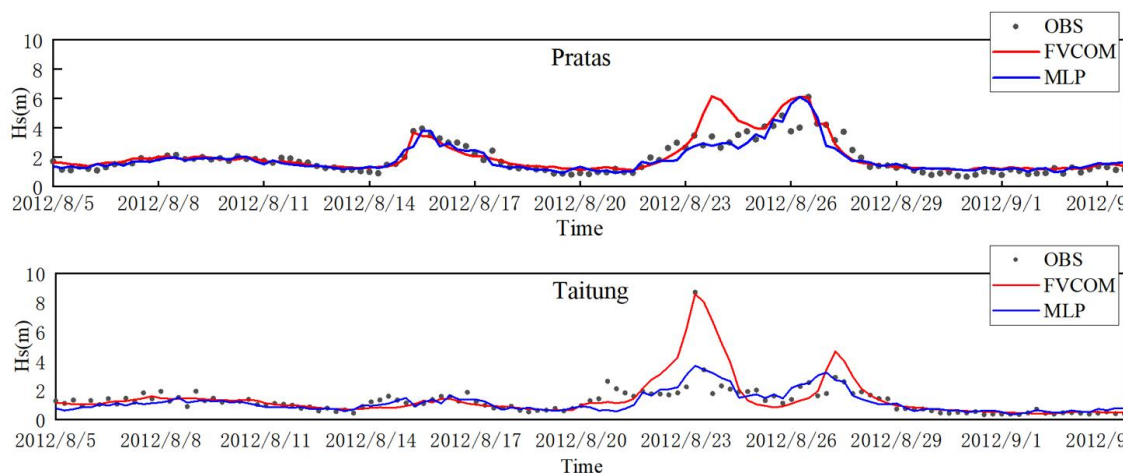


图 1.3 FVCOM-SWAVE 模型和 MLP 模型预测的有效波高与实测有效波高对比

同时，进一步比较 MLP 神经网络模型结果与数值模拟结果的空间平均波高情况。它们的交叉验证散点图如图 1.4 所示，其中 R^2 值为 0.82，RMSE 为 0.2m，发现 MLP 神经网络模型得到的平均波浪场分布和采用数值模式预报的平均波浪场具有相似性。

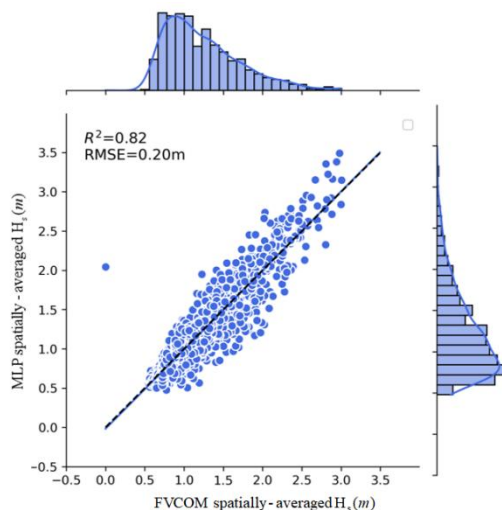


图 1.4 FVCOM-SWAVE 模型和 MLP 模型预测的 2012 年中国南海空间平均有效波高交叉验证散点图（黑色点划线代表 1:1 线，蓝实线代表最佳拟合线）

软件采用该 MLP 神经网络模型对 2013-2020 年有效波高进行预测，为验证预测结果的可靠性，图 1.5 为空间平均有效波高交叉验

证图 MLP 神经网络模型预测结果与基于 FVCOM-SWAVE 模型计算结果之间的关系，在 2013-2020 年， R^2 在 0.84 到 0.92，RMSE 在 0.21 到 0.16 之间，预测结果与数值模型计算结果有较好的一致性。同时，采用训练好的 MLP 神经网络模型对一年的有效波高相较于 FVCOM-SWAVE 模型进行数值模拟计算要更加快速。

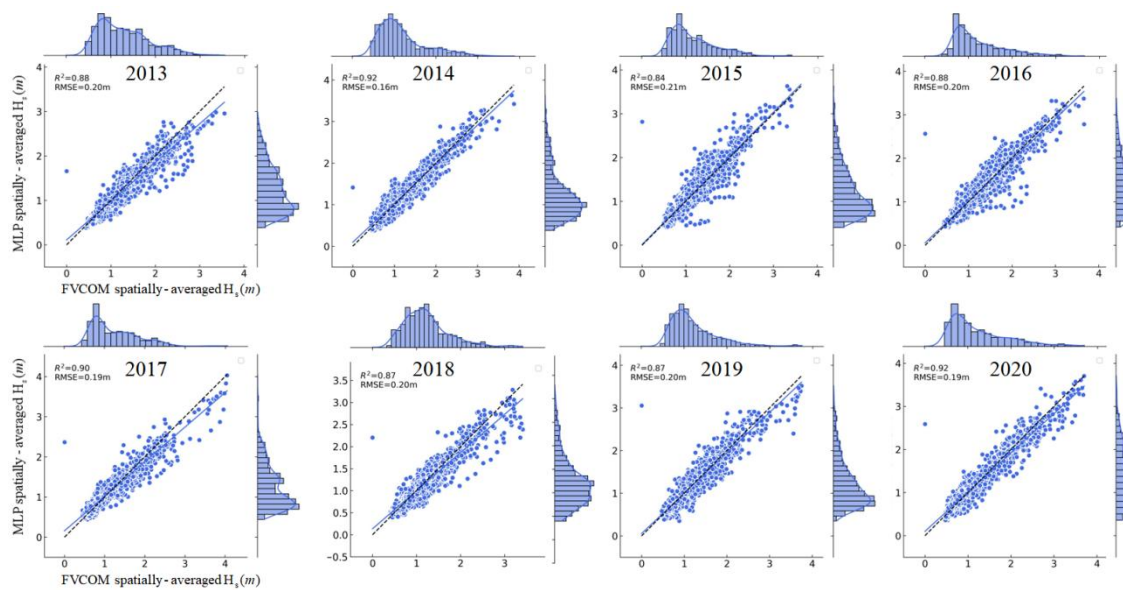


图 1.5 FVCOM-SWAVE 模型和 MLP 模型预测的 2013-2020 年中国南海空间平均有效波高交叉验证散点图（黑色点划线代表 1:1 线，蓝实线代表最佳拟合线）

综上所述，所建立海浪波高智能快速预报模型可以很好的预测中国南海的有效波高分布情况，结果相对可靠且准确。