

台风路径与强度预测

项目简介

我们主要在台风强度预测方面，应用一种基于空间注意力机制的融合网络(Spatial Attention Based Fusing Network, SAF-Net)模型。该模型将台风强度分解为 U Wind 变量和 V Wind 变量，并对每一个等压面引入空间注意力机制以提取风速分量的高响应区域，然后通过卷积融合对两种风速分量进行融合，最终实现通过学习各风速分量来预测合速度的 H 标。实验表明，使用 SAF-Net 模型可提高 24 小时台风强度预测的精度。

我们设计实现了一个基于 Flask 框架的台风路径与强度可视化系统，通过 Bootstrap 框架实现前端响应式布局、利用异步通信机制完成前端页面联动和后端数据更新，实现了台风路径和强度预报结果的可视化。

数据集

我们基于深度学习的台风路径与强度预测方法，主要使用了中国国家气象局(China Meteorological Administration, CMA)最佳路径数据集和欧洲中期天气预报中心(European Centre for Medium-Range Weather Forecasts, ECMWF)的再分析数据。将根据这两种数据预测西北太平洋区域台风未来。

预测相关方法

(1) CLIPER 方法

CLIPER 方法是台风路径与强度预测的最通用的方法，其本质是一个多元线性回归模型，该模型主要通过构建台风的气候性因子和持续性因子的特征，拟合 CLIPER 预测方程并预测未来的台风中心经纬度与风速。其表达式如下： $y = w^T x + b$

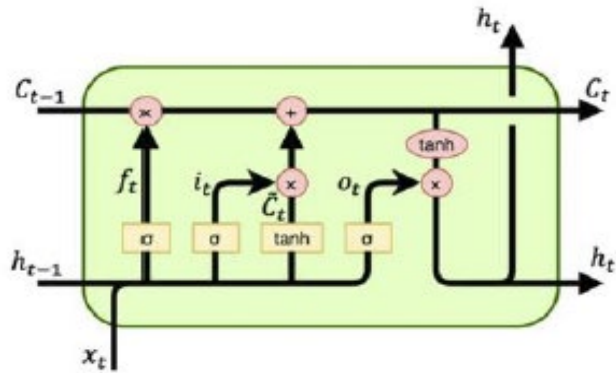
其中， y 表示模型预测值， $x = [x_1, x_2, \dots, x_n]$ 表示气候性因子和持续性因子， n 为因子总数， $w = [w_1, w_2, \dots, w_n]$ 为回归模型参数， b 为模型的偏置。

气候因子是指由再分析数据所刻画的大气环境中的各种影响气候变化的因子如：风速、气温、湿度等的信息。由于再分析数据维度高，往往需要降维处理后才作为 CLIPER 方法的因子；持续因子是指台风作为一个二维平面中曲线，其轨迹具有一定的运动规律性。该部分特征一般根据台风的中心经度、中心纬度、中心风速和中心气压等基础描述信息，以及常用物理与统计变量构建出的因子。

(2) RNN 模型

深度学习是一种具有强大表征能力的方法，随着深度学习技术的不断发展，目前多种深度学习框架已经在多个领域取得突破性发展，其中 RNN 便是其中一种重要的深度学习框架。RNN 是一种以序列数据为输入，并不断地在序列的演进方向进行递归，且所有神经单元链式连接在一起的神经网络。

自 Hochreiter 于 1991 年提出长短期记忆网络 (Long Short-Term Memory networks, LSTM) 解决 RNN 模型无法学习长时间序列信息的长期依赖问题后, RNN 模型逐渐走向成熟。LSTM 针对长期依赖问题新增了状态单元 C , 该单元只受当前时刻输入和上一时刻的输出影响, 因此长期记忆信息得到改善。此外, 新增的遗忘门 f 控制了上一时刻输出和记忆单元对当前单元的影响程度, 输入门 i 则控制了当前输入对状态单元的影响程度, 输出门 h 则控制当前输出内容的输入比例。LSTM 原理图如图所示。



其中, 遗忘门计算公式如下:

$$f_t = \sigma(U_f x_t + W_f h_{t-1} + V_f c_{t-1})$$

输入门计算公式如下:

$$i_t = \sigma(U_i x_t + W_i h_{t-1} + V_i c_{t-1})$$

当前记忆单元计算公式如下:

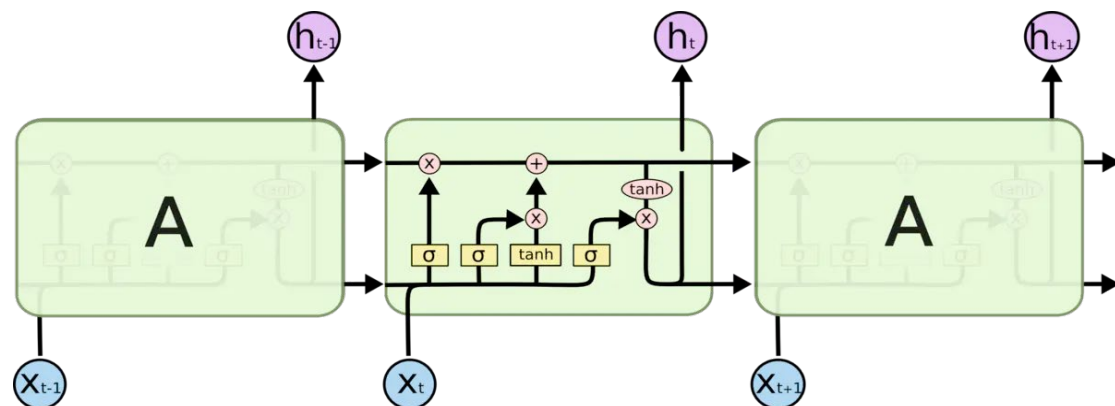
$$\tilde{c}_t = \tanh(U_c x_t + W_c h_{t-1}) \quad c_t = i_t \odot \tilde{c}_t + f_t \odot c_{t-1}$$

输出门的计算公式如下:

$$o_t = \sigma(U_o x_t + W_o h_{t-1} + V_o c_{t-1})$$

$$h_t = o_t \odot \tanh(c_t) \quad \text{式中 } \odot \text{ 表示哈达玛乘积。}$$

标准的 RNN 网络模型:



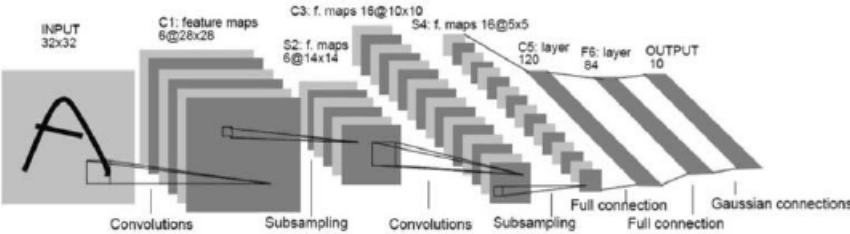
(3) CNN 模型

卷积神经网络 (Convolutional Neural Networks, CNN) 是一种以卷积计算为核心的前馈神经网络, 虽然 CNN 一般用于解决计算机视觉的问题。而大气再分析数据是描述地球三维

空间大气环境的重要数据，其等压平面构成了地球的三维空间表征，若将这些等压平面看作是描述大气的特征图，则可以使用 CNN 模型对其进行处理。一般 CNN 分为 2D CNN 和 3D CNN 两种卷积神经网络。2D CNN 最早于 1989 年由 LeCun Yann 提出，目前大多数研究都基于该结构，其具体模型如下图。LeNet 包含了深度学习的基本网络层：卷积层、池化层、全连接层等。其中卷积层是该网络的核心单元，该单元通过不同初始设置不同形状的卷积核，对输入的图像在不同位置做卷积操作，依靠监督学习的方式提取图像的特征。卷积操作具体公式表示如下：

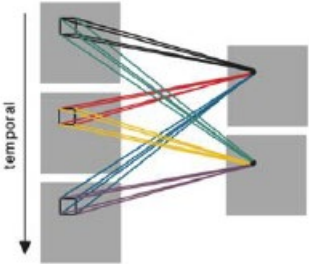
$$H(i, j) = (X * W)(i, j) + b = \sum_{k=1}^n (X_k * W_k)(i, j) + b$$

其中， $H(i, j)$ 表示对应图像矩阵的第 i 行第 j 列的输出结果； $*$ 代表卷积操作； X, W 分别表示输入图像矩阵和卷积核； b 表示为偏置项； X_k, W_k 分别表示输入图像矩阵和卷积核矩阵的第 K 维图像矩阵和卷积核矩阵。



LeNet 原理图

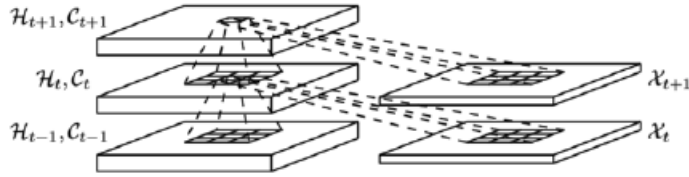
虽然 2D CNN 可以很好地提取图像的二维特征，但是客观世界却是三维空间，2D CNN 难以提取三维立体图像的特征。为了克服三维特征提取的问题 Ji 等人提出了 3D CNN，该网络结构与 2D CNN 的不同在于：2D CNN 的二维卷积核之间相互独立并不相互关联，而 3D CNN 的三维卷积核由多个二维卷积核构成，且这些卷积核之间相互关联，共同提取空间上的垂直信息，达到三维立体卷积的效果，原理图如下：



(3) ConvLSTM 模型

LSTM 在时间序列学习中取了良好的效果，但其在涉及时空问题上的效果却并不理想。虽然 CNN 可以通过采用 2D CNN 或 3D CNN 可以完成对二维或三维空间的表征，但却没有很好考虑时间维度上的空间信息表征。Shi 等人在保证提取时间特征的基础上，在 LSTM 单元加入用于提取空间特征的卷积运算，提出了 ConvLSTM 模型并在时空数据学习中取得了非常好的效果。

与 LSTM 模型相比，ConvLSTM 将模型的所有输入、输出、隐含状态以及各个控制门都改成了三维张量，以方便模型通过卷积操作表征空间信息，而其余的计算方法与 LSTM 相同。ConvLSTM 的内部结构如图所示。



其中，遗忘门计算公式如下：

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \circ C_{t-1} + b_f)$$

输入门计算公式如下：

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \circ C_{t-1} + b_i)$$

当前记忆单元计算公式如下：

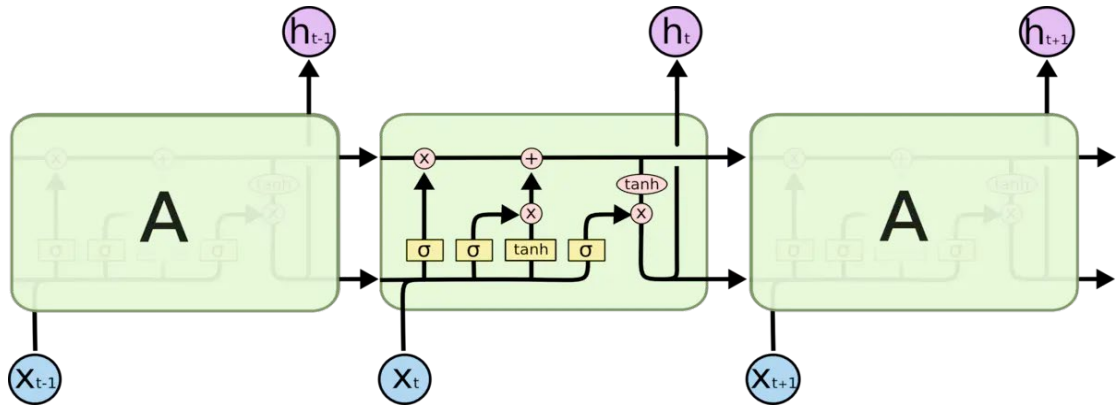
$$C_t = f_t \circ C_{t-1} + i_t \circ \tanh(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c)$$

输出门计算公式如下：

$$o_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + W_{co} \circ C_t + b_o) \quad H_t = o_t \circ \tanh(C_t)$$

式中*表示卷积运算； \circ 表示哈达玛乘积。

LSMT 模型：



相关评测指标

在台风路径预测任务中，使用了球面上真实经纬度点与预测经纬度点的两点距离误差指标来评测预测结果。在台风强度预测任务中，通过比较预测风速与真实风速之间的绝对误差来评价预测结果。

(1) 台风路径预测相关指标

使用球面距离误差公式表示预测经纬度点与真实经纬度点之间的差异：

$$D = 2R \sin^{-1} \left(\sqrt{\sin^2 \left(\frac{|a_p - a_r|}{2} \right) + \cos a_p \cos a_r \sin^2 \left(\frac{|g_p - g_r|}{2} \right)} \right)$$

其中， R 表示地球半径一般取值为 6371.009km， a_p ， a_r 分别表示预测纬度和真实纬度， g_p ， g_r 分别表示预测经度和真实经度。

(2) 台风强度预测相关指标

使用均绝对方误差 (Mean Absolute Error, MAE) 来表示预测风速与真实风速之间差异：

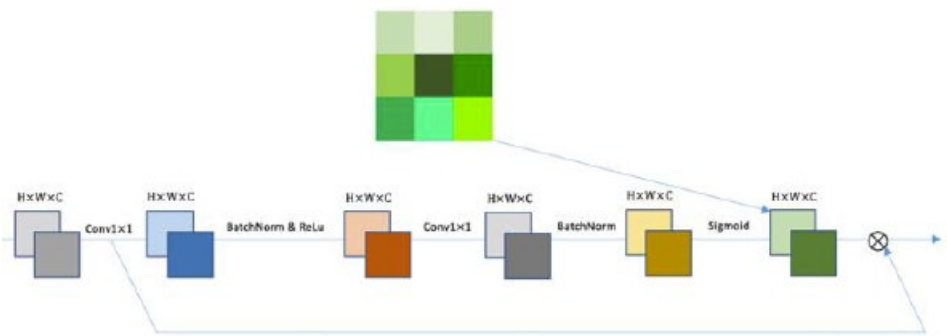
$$MAE = \frac{1}{n} \sum_{i=0}^n |w_p - w_r|$$

其中， w_r 表示真实风速， w_p 表示预测风速。

SAF-Net 中的空间注意力机制

台风强度预测问题与台风路径预测问题的三维结构主要区别是：第一，台风的路径是一种具有方向的矢量，而台风的强度不仅是一种具有方向的矢量，其更包含 U Wind 与 V Wind 两个速度分量的物理含义；第二，台风路径需要使用 Geopotential 描述的台风内部能量结构和地势高度，而台风的强度则需要使用 U wind 和 V wind 共同描述；第三，台风的路径是一个空间大尺度预测问题，可以根据通道注意力机制进行学习其移动方向，而台风的强度是空间多尺度问题，可以根据空间注意力机制进行学习其力分布情况。空间域的注意力机制是将注意力集中在同一个特征图不同区域，以选择出每一个特征图的重要区域。

为学习不同等压面下 U wind 与 V wind 的注意力值，我们引入了另一种空间注意力机制。该注意力机制与 CBAM 相类似，其原理如下：首先，对输入的不同等压面做一次 1×1 的卷积操作，相当于对所有等压面做了一次非线性变换，得到结果后进行批归一化和 ReLU 非线性变换；然后，对结果再做一次 1×1 的卷积操作，得到结果后进行最后一次批归一化，通过 Sigmoid 函数得到注意力值；最后，利用得到的注意力值与原等压面进行哈达玛乘积，得到经过注意力机制后的不同等压面的风速。原理如图：



SAF-Net 中的融合机制

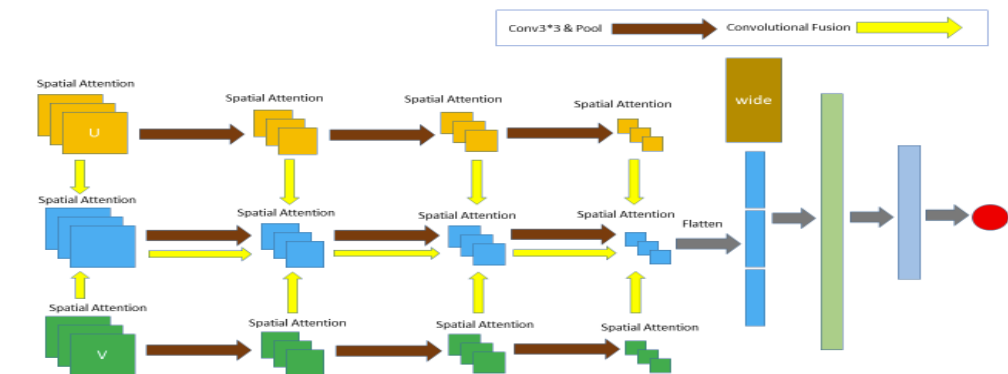
由于采用 U Wind 与 V Wind 两个分速度分别建模的方法，因此需要采用适当的融合策略。在机器学习和深度学习领域，一般通过特征融合的方法进行融合。

特征融合是指在模型的中间层，根据不同的特征特点进行数学组合，从而得到更好表达的特征的过程。特征融合技术大致可以分为两种：基于机器学习的特征融合技术与基于深度学习的特征融合技术。

基于机器学习的特征融合技术主要包括：多个分类器的贝叶斯决策，该方法通过先验概率计算出后验概率，为不同的分类器计算融合权重，最后实现加性融合；利用稀疏表示矩阵进行特征融合，该方法通过对样本提取多种特征并建立特征联合稀疏表示矩阵，利用该矩阵完成多维特征选择，常见方法如 Lasso 回归。

基于深度学习的特征融合技术一般在隐藏层中实现特征融合，主要包括空间特征融合技术和时间特征融合技术。空间特征融合技术主要的融合对象是 CNN 的输出特征图，该方法通过对特征图进行一系列数学组合运算，完成特征的融合表达；时间特征融合技术主要的研究对象是 RNN 的输出序列结构或含时间信息的 CNN 输出的特征图，该方法通过对不同的时间序列输出进行池化，以完成特征融合。

基于 wide&deep 异构模态台风强度特征融合模型



对于 wide 部分，根据 CMA 数据集构建了台风强度二维时序特征，并采用 GLM 作为 wide 部分模型。对于 deep 部分，根据 U Wind 与 V Wind 变量，构建台风及其周边的三维时序结构。本文提出的 SAF-Net 用于处理台风三维结构。SAF-Net 的中间蓝色部分网络为主干网络，分别标示有 U 的橙色部分和 V 的绿色部分的网络为枝干网络，且两个枝干网络结构相同，均用于提取台风分速度的特征。首先，分别对输入的 U Wind 和 V Wind 变量做一次空间注意力，得到两个输出后在主干网络进行卷积融合，得到特征图 R^1 ；接着，两个枝干网络同时做一次 3×3 的卷积操作和最大池化操作，所得结果进行空间注意力并在通道维度上进行融合后，得到特征图 R^2 ；然后，主干网络的特征图 R^1 与枝干网络进行相同的操作得到特征图 R^2 ，将 R^2 与 R^2 再进行一次卷积融合，得到主干网络的最新输出特征图 R^2 ；最后，枝干网络与主干网络重复第二步操作两次后，将最终结果做一次 Flatten，完成 deep 部分的表征过程。得到 wide 部分与 deep 部分表征后，采用联合训练 (jointly train) 的方式，并增加一层神经网络融合两部分信息，最后得到台风 24 小时后强度的预测值。计算模型训练的损失，使用平均绝对误差 (Mean Absolute Error, MAE)，计算预测的强度 (\hat{y}) 与真实强度 (y) 之间的误差值：

$$\text{Loss} = \frac{1}{n} \sum_{i=0}^n |y - \hat{y}|$$

人机交互界面

人机交互界面：提供给用户可视化操作界面，用户可以在页面选择相应条件，在地图上

生成历史台风信息、预测台风信息。用户可查看各台风记录的具体信息和台风预测报告。系统前端使用 HTML5、jQuery、CSS3、Bootstrap 开发完成，后端使用 Flask、Python 开发完成。

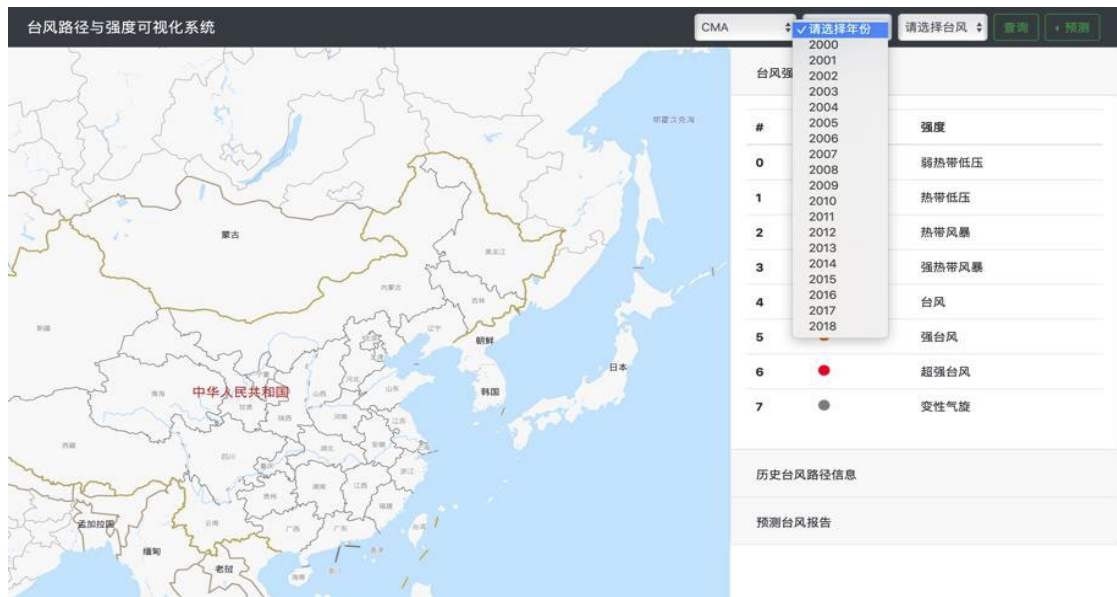
历史台风路径与强度可视化模块需要用户在 Web 系统提供两个输入：台风发生的年份、所选年份台风名称。系统得到用户输入后，具体处理过程如下：可视层利用 Ajax 技术将用户输入传递到逻辑层；逻辑层根据输入，利用 Pandas 包连接数据层，访问 CMA 数据集并通过数据筛选函数取出符合用户选择条件的所有台风记录，所选择记录按时间顺序排列返回到逻辑层；逻辑层根据数据层返回的结果以 Ajax 方式回调输出；可视层利用 Ajax 请求成功后的回调函数获取逻辑层的返回结果，并将返回的字符串包装成用于显示的台风记录，其内容包括：台风发生时间、台风中心经度、台风中心纬度、台风中心风速、台风中心气压、台风强度。最后利用百度地图的 JS API 根据所提供的经纬度、并使用 jQuery 以动态的方式将台风路径显示在百度地图上，且用不同颜色的圆点标示台风强度的强弱。

预测台风路径与强度可视化模块与上面历史台风路径与强度可视化模块类似，就不赘述。

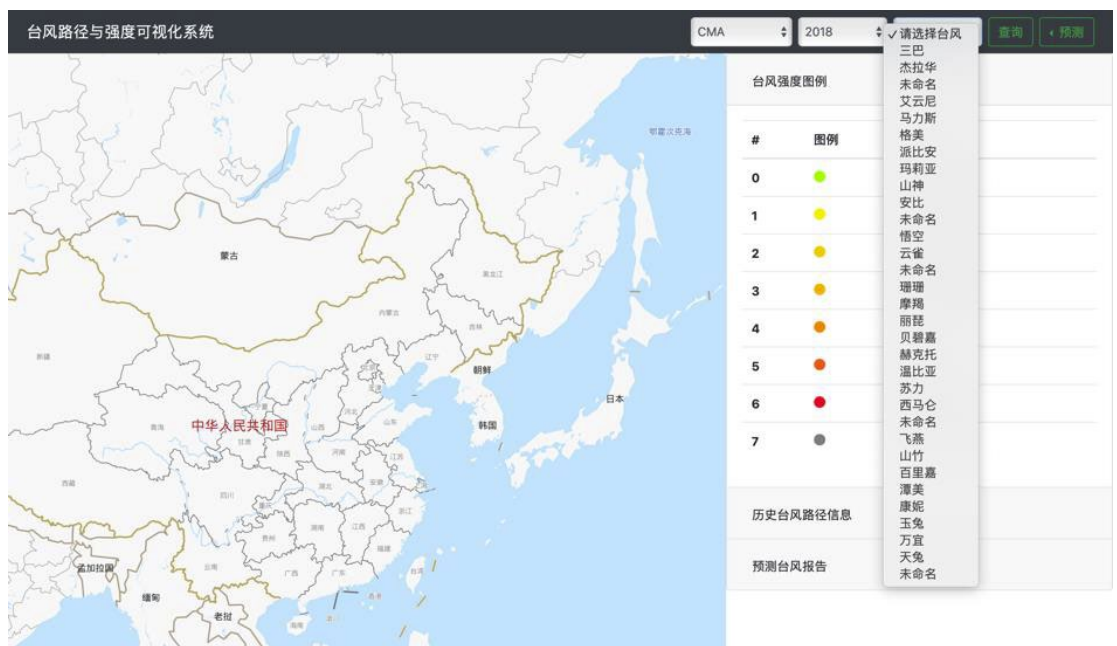
将 Web 系统部署到服务器后，用户可以通过 IP+端口号的方式访问到台风路径强度可视化系统，本地部署可通过 127.0.0.1:8080 访问。网页首页如图所示。网站首页提供台风发生年份下拉选择、台风名下拉选择，并有查询和预测按钮。页面左侧为基千百度地图的台风路径与强度展示地图，右侧为台风强度图例、历史台风路径信息、预测台风报告。



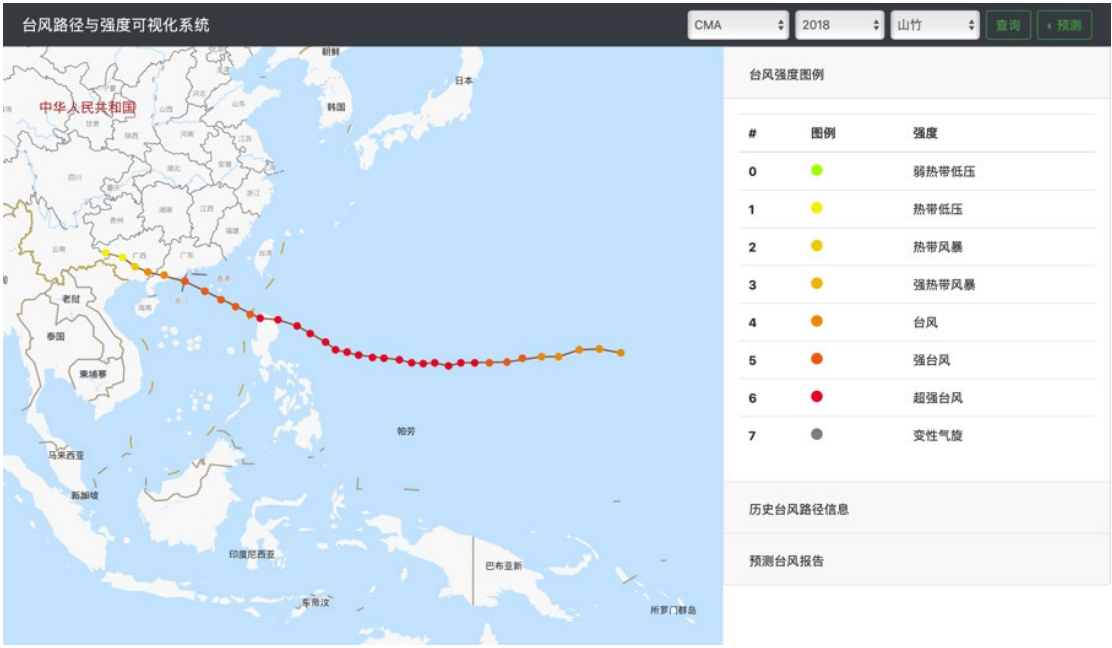
用户点击“请选择年份”可产生年份下拉列表，目前可选年份为 2000 年至 2018 年，如图：



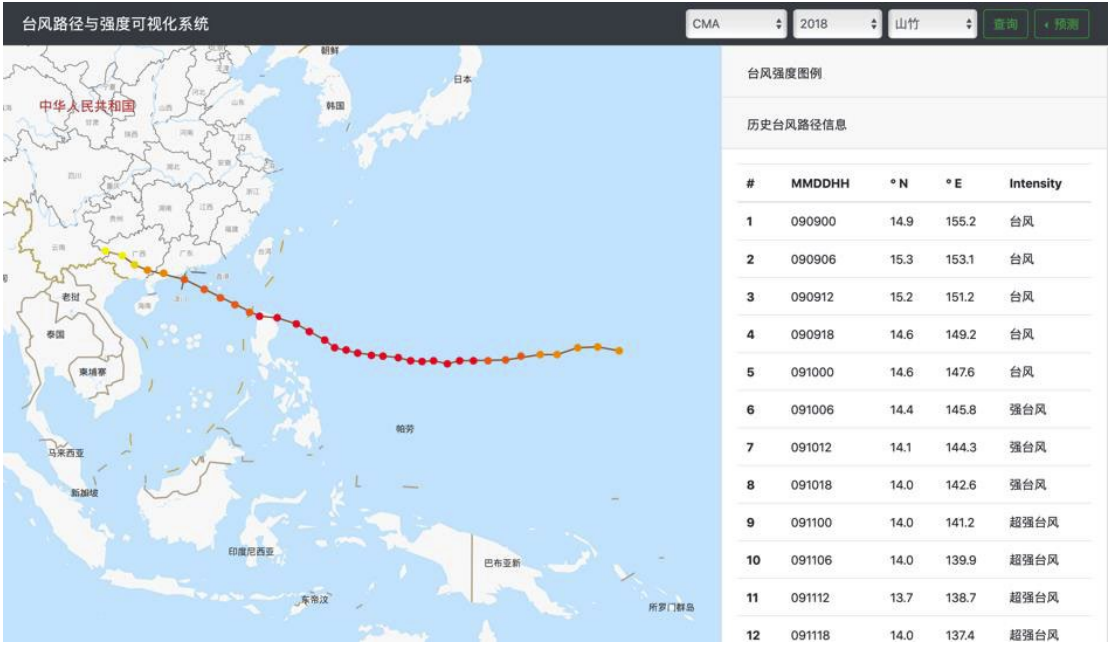
用户点击“请选择台风”可得到所选年份的台风名，如图：



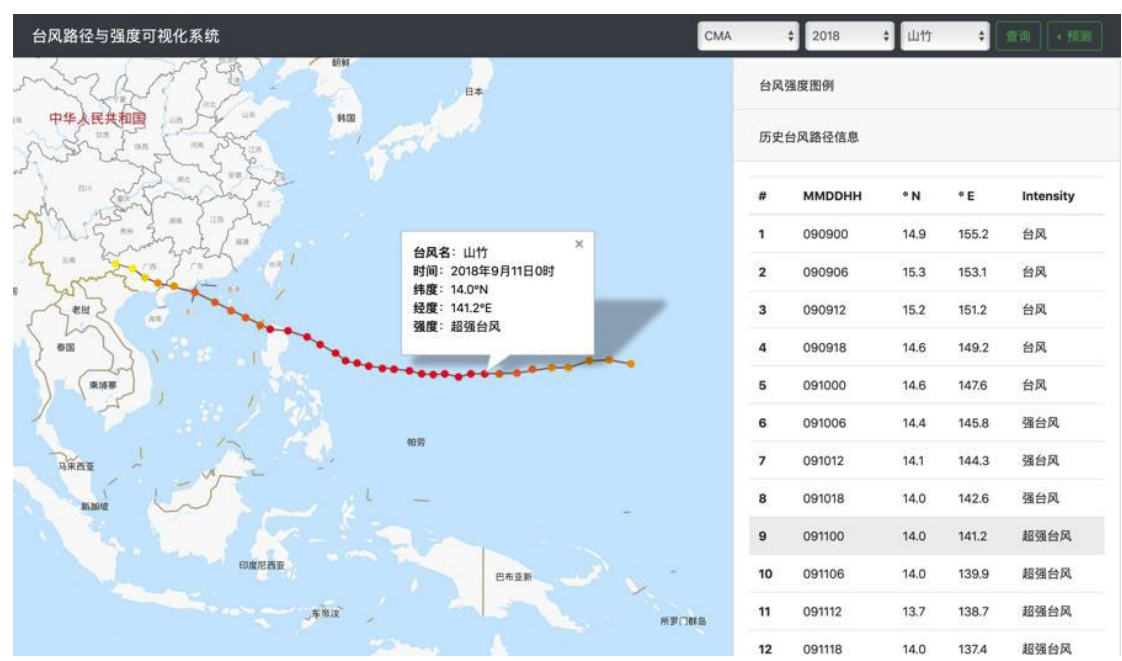
用户点击“查询”，Web 系统将动态生成台风路径并展示在左侧地图上，如图：



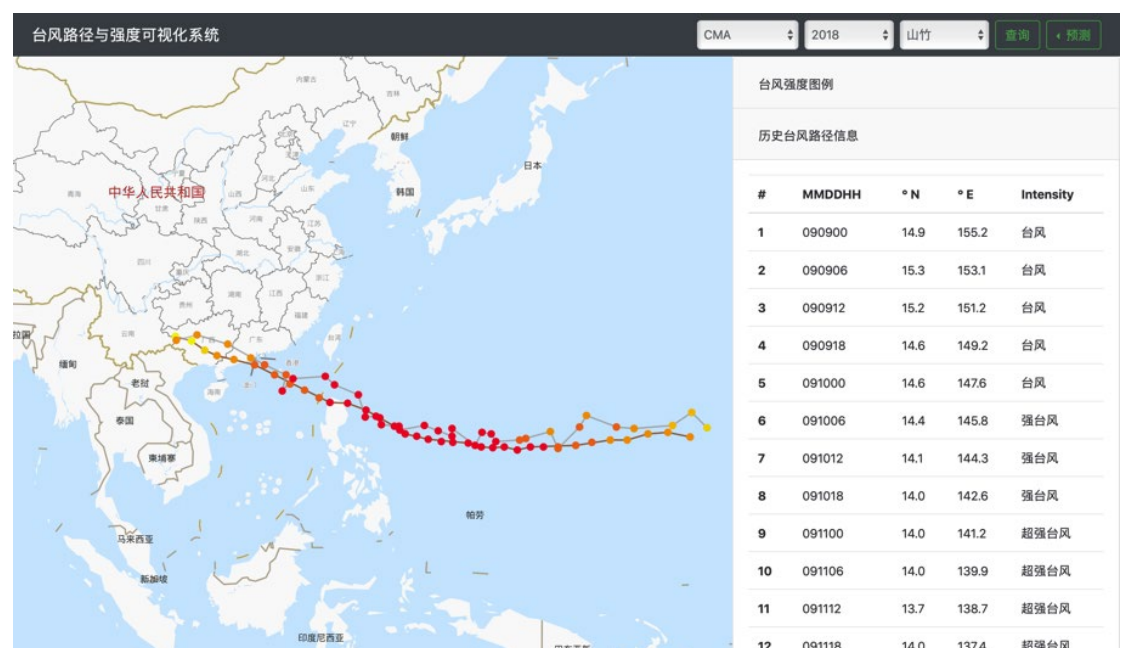
用户点击右侧下拉栏中的“历史台风路径信息”，查看台风的历史信息，如图：



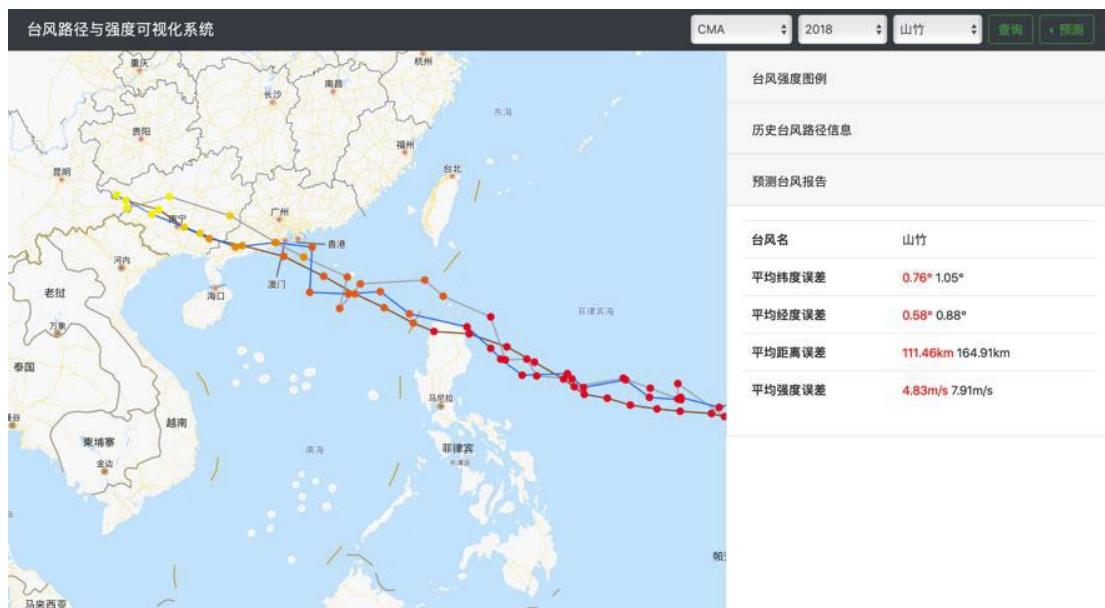
用户点击右侧栏台风记录表的每一行或直接点击左侧地图的台风中心，可查看具体时间台风的具体信息。如图：



用户点击“预测”，选择“基线模型”，可以显示基线模型（灰色线）的预测结果。如图：



用户点击“预测”，选择“我的模型”，得到提出模型（深蓝色线）的预测结果，点击“预测台风报告”可以对比本文提出模型（红色）与基线模型（黑色）对比结果。如图：



项目分工

1953729 吴浩泽：负责路径预测、项目文档

1952211 黄金坤：负责的强度预测、PPT 答辩和项目演示

共同完成：flask 前端演示项目