



浙江水利水电学院

Zhejiang University of Water Resources and Electric Power

本科毕业设计（论文）

（2021 届）

题目：基于神经网络的小流域山洪预警技术研究

学 院	信息工程学院
专 业	软件工程
学 号	2017b11081
学生姓名	王凯
指导教师	汪松松
提交日期	2021 年 5 月 28 日

声明及论文使用的授权

本人郑重声明所呈交的论文是我个人在导师的指导下独立完成的。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写的研究成果。

论文作者签名：

年 月 日

本人同意浙江水利水电学院有关保留使用学位论文的规定，即：学校有权保留送交论文的复印件，允许论文被查阅和借阅；学校可以上网公布全部内容，可以采用影印、缩印或其他复制手段保存论文。

论文作者签名：

年 月 日

摘 要

浙江省总体地形西高东低，西部多丘陵山地地形，山区内部小溪小河较多，夏季降雨频繁，并且持续性暴雨较多，给山洪暴发制造了有利条件，人民的生命财产安全受到了严重威胁。国内外的山洪防治往往采用动态雨量，水文信息处理，水动力模型等措施进行预测，使用单一的降雨量数据对水位进行预测，该方法对于大江大河预测效果较好，但对于小流域山洪发生具有突然性的地区预测效果较差。

本文以浙江临安地区苕溪的桥东村站点作为研究区域，对山洪灾害水位预测进行研究。基于神经网络可以输入多维的数据优势，本文从浙江实时雨水情网站获取水位和降水历史与实时信息，从 tutiempo 网站获取湿度以及气温的历史信息，作为神经网络模型的输入。本文设计基于时间序列的 RNN 和 LSTM 对比预测模型，使用 python3.7 和 tensorflow2.0 框架开发，在 jupyter notebook 平台编写 RNN 神经网络以及 LSTM 神经网络核心代码。通过采用不同的模型参数集，对两种模型进行训练。最后对 RNN 神经网络以及 LSTM 神经网络的拟合情况进行对比，并对连续 48 小时的水位预测情况进行对比研究。当模型预警的水位线超过桥东村预警水位线时可以发布预警信息，从而降低山洪灾害的损失。

实验表明：（1）RNN 神经网络和 LSTM 神经网络在最优参数的情况下，模型训练均方误差分别为 0.000108 和 0.000064，对近期水位进行预测均方误差分别为 0.003473 和 0.001545。（2）在足够的训练次数和合适的参数情况下 LSTM 与 RNN 相近，综合不同的模型来看，LSTM 神经网络的预测效果好于 RNN 神经网络。

关键词：山洪预警；小流域；LSTM；RNN

Abstract

The overall terrain of Zhejiang Province is high in the West and low in the East. There are many hills and mountains in the West. There are many streams and rivers in the mountains. There are frequent rainfall in summer, and there are many persistent rainstorms, which create favorable conditions for the outbreak of mountain torrents and seriously threaten the safety of people's lives and property. At home and abroad, dynamic rainfall, hydrological information processing, hydrodynamic model and other measures are often used to predict the mountain flood control, and single rainfall data is used to predict the water level. This method has good prediction effect for large rivers, but poor prediction effect for areas with sudden mountain flood in small basins.

Taking Qiaodong Village Station of Tiaoxi in Lin'an area of Zhejiang Province as the research area, this paper studies the prediction of mountain flood disaster water level. Based on the advantage that neural network can input multi-dimensional data, this paper obtains the historical and real-time information of water level and precipitation from Zhejiang real-time rainfall website, and obtains the historical information of humidity and temperature from tutiempo website as the input of neural network model. In this paper, a time series based RNN and LSTM comparative prediction model is designed, which is developed by using Python 3.7 and tensorflow 2.0 framework. The core code of RNN neural network and LSTM neural network is written on jupyter notebook platform. By using different model parameter sets, the two models are trained. Finally, the fitting of RNN neural network and LSTM neural network is compared, and the prediction of water level for 48 hours is compared. When the model warning water level exceeds the Qiaodong village warning water level, the warning information can be released, so as to reduce the loss of flash flood disaster.

The experimental results show that: (1) under the condition of optimal parameters, the mean square error of RNN neural network and LSTM neural network is 0.000108 and 0.000064 respectively, and the mean square error of short-term water level prediction is 0.003473 and 0.001545 respectively (2) In the case of enough training times and appropriate parameters, LSTM is similar to RNN. Considering different models, the prediction effect of LSTM neural network is better than RNN neural network.

Key Words: Mountain flood warning; Small watershed; LSTM;RNN

目 录

摘 要.....	I
ABSTRACT.....	II
第 1 章 绪论.....	1
1.1 选题的背景及意义.....	1
1.1.1 选题背景.....	1
1.1.2 选题意义.....	1
1.2 国内外研究现状.....	2
1.2.1 国内研究现状.....	2
1.2.2 国外研究现状.....	2
1.2.3 发展趋势.....	3
1.3 研究内容与步骤.....	4
1.3.1 主要研究内容.....	4
1.3.2 研究步骤.....	4
第 2 章 小流域山洪预警模型分析.....	5
2.1 苕溪小流域特征分析.....	5
2.2 RNN 模型山洪预警分析.....	5
2.2.1 模型选择.....	5
2.2.2 RNN 神经网络的分析.....	5
2.2.3 RNN 神经网络的局限性.....	7
2.3 LSTM 模型山洪预警分析.....	7
2.3.1 LSTM 神经网络的原理.....	8
2.3.2 LSTM 模型山洪预警的优势.....	9
2.4 本章小结.....	9
第 3 章 小流域山洪预警模型设计.....	10
3.1 模型输入输出参数确定.....	10

3.1.1 预警指标输入参数确定	10
3.1.2 预警信息输出参数确定	11
3.2 时间序列预测	12
3.3 模型实现流程	12
3.4 本章总结	14
第 4 章 小流域模型的实现	15
4.1 模型的实现环境	15
4.2 数据的获取	15
4.2.1 水位和雨量信息获取	15
4.2.2 空气湿度温度信息获取	16
4.3 模型的代码的编写	16
4.3.1 数据读入以及归一化	16
4.3.2 构造神经网络模型	16
4.3.3 绘制预测图像	16
4.3.4 连续预测 48 小时的水位	17
4.4 水位的时间序列预测	17
4.4.1 RNN 神经网络的水位预测	17
4.4.2 LSTM 神经网络的水位预测	22
4.5 模型预测近期水位	26
4.5.1 RNN 神经网络预测近期水位	26
4.5.2 LSTM 神经网络预测近期水位	28
4.6 结论与应用	31
4.6.1 模型对比及结论	31
4.6.2 模型预测实战	32
4.7 本章小结	32
结论	33
参考文献	34
致谢	35

第1章 绪论

1.1 选题的背景及意义

1.1.1 选题背景

浙江省位于我国的东部丘陵地区，西高东低，西部山地丘陵较多，且是许多河流的发源地，东部较为平坦^[1]。而且气候受季风环流影响显著，夏季高温多雨，冬季低温少雨，春夏之间，锋面气旋流动频繁，雨量增多，受其影响，大雨暴雨天气屡次发生。此外，浙江省丘陵地形和由于过度砍伐导致自然环境破坏严重，进而山洪、泥石流等地质灾害不断发生^[2]。山洪不仅对居民财产和生活情况造成影响，极大地制约了浙江丘陵地区的社会经济增长。我国有超过 70% 的灾害统计是由山洪或者中小流域河流泛滥所导致的。山洪、中小河流流域等自然灾害是防洪救灾工作中不可忽视的环节。近年来，随着全球气候变暖和厄尔尼诺现象等极端天气事件的频繁发生，我国大部分地区特别是南方地区自 2020 年汛期以来，经历了多个强降水周期。据水利部统计，截至 6 月底，已有 16 个省的 198 条河流被洪水淹没在警戒线以上。至 7 月 9 日 14:00，洪水灾害影响了 3020 万人在 27 个省(自治区、市)，对农作物造成巨大影响，并且产生了巨大的经济损失^[3-4]。

1.1.2 选题意义

传统水位预测方法大多是水位数据分析法，构建水位模型等，需要大量的数据，对于小流域的数据较难实现。近年来，人工智能技术方兴未艾，其中神经网络广泛应用于信息处理、系统识别、智能检测、汽车工程、化学工程、卫生保健、人脸识别等各个方面。如果将该技术应用到山洪灾害防治领域，可以帮助处理中小流域更加准确的进行山洪预警^[5]。由于神经网络数据量需求小，准确性高等特点，对于小流域的水位预测具有天然的优势。近年来 LSTM 神经网络，GRU 神经网络开始广泛得到应用，并且衍生出了如 LSTM-KNN 等优秀的神经网络，更好的使用神经网络对于山洪灾害的预测对于我国山洪灾害的防治具有重要的意义。

综上所述，研究构建基于神经网络的小流域山洪预警模型是应对小流域山洪灾害的有效措施。

1.2 国内外研究现状

1.2.1 国内研究现状

山洪灾害是我国近 20 年来的一个热点。虽然我国山洪灾害预防起步较晚，但是提升较快^[6]。依赖于网络系统的国家防汛抗旱指挥中心，在中国 31 个省和七大流域基本上实现了水文预警系统的建设，提高了我国各地山洪预警的时效性。因为水文网的全面建设为山区洪水预警系统奠定了坚实的基础。随着自动化，随着物联网、传感器技术等先进手段的快速发展，大量先进的水文监测设备应用于防灾预警，实现了山洪监测、水文监测、地理监测的实时自动监测。使得水文信息获取更加快速，自动，并且可以远程传送，这使得山洪预警更加及时、准确。当今，我国大量学者对于临界雨量的分析计算主要使用经验方法和理论方法两种^[7]。经验方法主要根据实验的经验，以及大量数据的相关性确定临界雨量的指标，并无相应的物理机理。理论方法以水文学、物理学和地质学为基础，它对数据要求高。物理机制和推导过程严谨，主要有水位反推法、土壤饱和-雨量关系法、雨量临界曲线法等。

山洪灾害的形成机理十分复杂。土壤水分、植被、山体坡度、雨量等诸多因素与山洪灾害的形成密切相关。雨量是引发山洪灾害的关键因素^[8]。当某一小流域某一时段的雨量达到或超过某一临界值时，形成的洪水超过了河流的泄洪能力，此时的雨量可能引发山洪灾害，称为临界雨量。在 2020 年由刘莫阳^[9]发表的关于 LSTM-KNN 模型在中国不同实现洪水预测的适用性。事实证明了 LSTM-KNN 模型在洪水预报领域的有效性，并且强调了 LSTM 用于真实洪水预测的潜力。

1.2.2 国外研究现状

在阿尔及利亚洪水是该国最严重的自然灾害之一^[10]。Boudani Mustapha Rabie 发现，由于在当前全球气候变化的背景下，山洪灾害的频率和幅度在空间和时间上变得越来越不规律，因此促使我们改进管理和预测战略。Boudani Mustapha Rabie 希望为阿尔及利亚东北部地区建立一个使用最少数据来进行洪水预报和管理的系统，该模型是在 Sokolovsky 综合水文模型的基础上，结合 Horton 流域离散化体系结构，通过分析能够适应实时收集的气候数据的水文数据，来预测水文网络所有部分的短期洪水过程。他在过去的降雨灾害模拟中获得了峰值流量振幅 65.2% 的准确度和地表径流基准时间的 88.3% 左右的精度。这种低成本的简单模型为洪水管理提供了更多的可能性，并且可以通过更好的空间化和更多的现场数据校准来改进。

在 2020 年，加拿大的科学家 M. Esfandiari^[11]使用随机森林的方法模拟 Fredericton 在 2018 年的山洪，并且分析了山洪不同的影响因子之间几种组合关系。根据洪峰前后日可用的 sentinel-2 光学卫星图像对这些影响因子进行测试。当仅使用海拔、坡度、坡向、距河距离、植被覆盖率五个因子时准确率是最高的，达到了 97.57%。

对于山洪的水流主要来自于积雪堆积和融化的寒冷地区，春汛洪水预报成了一个严峻的挑战^[12]，美国利用 grace 卫星资料建立了寒冷区域的洪水预报模型。该模型通过模拟融雪产生的地表径流以及相应的地下水流量来预测洪水。利用气温指数模型预测融雪区域的地表径流。该模型主要被应用在红河流域，预测效果取得 95% 的准确率，预测结果良好。结果表明，最大融雪量与观测站的流量峰值存在两周的滞后关系。虽然模型相对较简单，只需要输入温度和 grace 卫星资料，但经过简单的修改后很容易应用到其他地方，对于数据量较少的地区尤其游泳，该模型对比大多数传统的洪水预报方式拥有极大的优势。

Xuan-Hien Le 在 2019 年发表的关于 LSTM 神经网络对于洪水预警中的作用^[13]。选用了越南的 Da River basin 河作为研究对象并将 1985 以前的数据作为输入数据集组合用于流量预报，结果表明 The Nash-Sutcliffe efficiency 达到了 99%，95%和 87%表明了 LSTM 神经网络对于流域水位预测的可适用性。

1.2.3 发展趋势

神经网络从二十世纪五十年代开始就已经拉开了序幕，在 1960-1980 年都是较缓慢的发展，直到 20 世纪 80 年代提出了全连接的反馈神经网络以及误差反向传播神经网络，也就是 BP 神经网络，神经网络开始快速发展。21 世纪初硬件技术也开始快速成长，更加适合神经网络的计算，神经网络开始进入一个蓬勃发展期。如今，神经网络已经应用于各个领域，比如信号处理、智能检测、汽车工程、化学工程、卫生保健、人脸识别等各个领域。

神经网络在山洪预警的使用也越来越频繁，由于神经网络的灵活性强，并且使用者无需对该领域很深刻的理解也可以编写程序，取得相当好的成绩。

神经网络与专家系统的融合使神经网络可以通过大规模的简单信息来模拟人脑对信息的处理。另外深度神经网络到达一定层数后会出现梯度消失或者梯度爆炸等问题，将各种神经网络混合使用并且提出更加有效的神经网络将具有深远的影响。

1.3 研究内容与步骤

1.3.1 主要研究内容

通过对国内外研究现状以及发展趋势的分析，发现神经网络模型预测具有明显的优势。本文主要的研究对象是茗溪，研究的水文信息获取地点是桥东村站点，站点代码为 63000200。该地区位于南茗溪与南溪的交汇点，两面环山，是水流汇集之处，夏季降水较多时水位上涨较快。国内外对于山洪预警往往采用水文模型法，但水文模型法对于水文数据的要求较高，数据量要求较大，在小流域地区很难建立合理的水文模型。由于神经网络对于专业知识的依赖度低，准确度较高，因此使用神经网络对于山洪预警具有显著的优势。

本文将使用水位，雨量，空气湿度，温度作为输入参数通过循环神经网络（简称 RNN 神经网络）与长短期记忆模型循环神经网络（简称 LSTM 神经网络）构建模型，输出水位预测信息，并且对连续 48 小时的水位数据进行预测，从而达到山洪预警的目的。使用 python3.7 版本，tensorflow2.0 框架编写 RNN 神经网络与 LSTM 神经网络，对比分析在不同参数下的 RNN 神经网络与 LSTM 神经网络，并且使用各个模型预测连续 48 小时的水位信息，评测准确情况，选择出在该区域更加优秀的神经网络模型。

1.3.2 研究步骤

（1）收集国内外关于山洪预警方法与神经网络在山洪预警方面的文献资料，最常见的神经网络模型有 BP 神经网络，SVM，RNN 神经网络，LSTM 神经网络等。分析引发山洪的因素，选定将水位、降水量、空气湿度、温度作为输入特征，水位作为输出特征。分析各种神经网络的优劣性，选择 RNN 神经网络与 LSTM 神经网络作为研究模型。

（2）通过网络爬虫从浙江省雨水情监测预警平台获得雨量与水位信息，从 tutiempo 网站获取湿度以及气温的历史信息，对数据进行清洗，错误值使用平均值进行填充，共爬取 7296 小时的数据。

（3）使用 tensorflow2.0 编写 RNN 以及 LSTM 神经网络模型，将水位，雨量，温度，湿度作为输入特征，水位信息作为输出特征，通过调整模型参数来形成不同的模型，并且对不同参数下的模型进行对比分析，每个模型进行连续 48 小时的水位预测，对水位预测误差进行计算，分析模型的优劣。

第2章 小流域山洪预警模型分析

2.1 茗溪小流域特征分析

茗溪是浙江省西部的一段河流，本文将茗溪中南溪与南茗溪汇合地作为研究地点，水位站点为桥东村站，站点编号为 63000200。茗溪水流湍急，全年平均降雨量可达 1400mm，梅汛期降水量为 480mm；台汛期，降水量约为 250mm，汛期降水量为全年的 70%左右。桥东村站点位于南溪与南茗溪汇集点下游约 500m 处，该站点两面靠山，河流从两山之间穿过，汛期雨量较多时，山中溪水汇入茗溪，水位上涨较快，易发生山洪灾害。桥东村水位站的预警水位线是 84.66 米，保证水位线是 85.66 米。

2.2 RNN 模型山洪预警分析

2.2.1 模型选择

常见的神经网络预测模型有全连接神经网络，循环神经网络，卷积神经网络等。卷积神经网络常用于图像处理，并且对于拥有时间序列的水位，雨量等信息不适合使用。全连接神经网络常常搭配其他神经网络进行预测，对于时间，并且无法根据时间变化来记忆特征参数，从而对于时间序列的预测效果较差，因此不使用全连接神经网络。

RNN 神经网络的优势在于它的同一个隐藏层中前后两个神经元可以相互传递消息，使前一个输出会对后一个的输出产生影响，从而达到按照序列预测数据的目的。

本文主要是对水位数据进行预测，获取到的数据是按时间序列进行分布，因此 RNN 神经网络相比于全连接神经网络和 LSTM 神经网络具有很强的优势。

2.2.2 RNN 神经网络的分析

单个循环神经网络结构如图 2-1 所示， X_t 是一个输入数据，循环核表示记忆体， h_t 里面的内容由 x_t 与权重矩阵 whx 进行线性变换加上，上一个循环核所传送过来的内容即 whh 与 s_{t-1} 进行线性变换的结果。再过激活函数即为 y_t 的值。 y_t 示记忆体 h_t 与权重矩阵 why 进行线性变换，再过激活函数的值。

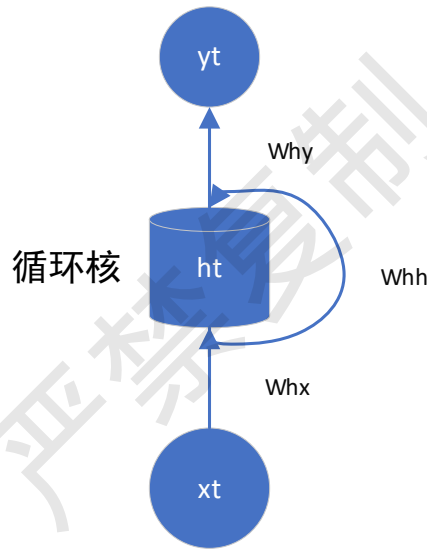


图 2-1 简单的循环神经网络

循环神经网络按照时间步展开，就是把循环核按照时间轴方向展开，记忆体 h_t 周围的权重矩阵在 w_{hy} 、 w_{hh} 、 w_{hx} 在前向传播过程中是不变的，训练完成后使用效果最好的权重矩阵 w_{hy} 、 w_{hh} 、 w_{hx} 执行前向传播，并且输出预测结果。只有当进行反向传播时才会更新参数矩阵。

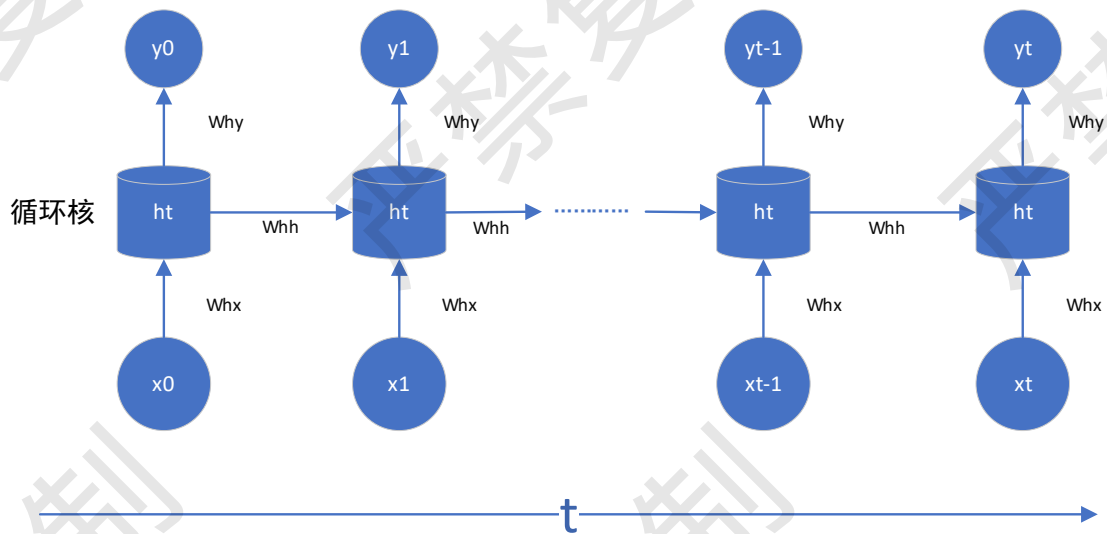


图 2-2 RNN 时间线展开图

神经元节点 y_t 通过公式 2.1, 2.2 来进行计算。网络的输入输出变化均随时间变化，因此我们称为序列。

$$h_t = \tanh(x_t * w_{hx} + h_{t-1} * w_{hh} + b_h) \quad (2.1)$$

$$y_t = \text{softmax}(h_t * w_{hy} + b_y) \quad (2.2)$$

2.2.3 RNN 神经网络的局限性

RNN 神经网络最大的特点就是存在循环核，循环核可以存储状态信息，正是因为每个时间步都存在状态信息，因此可以时间按照时间序列进行预测。但如果需要展开的循环时间步过长，会造成梯度消失的问题^[14]。RNN 梯度消失问题如图 2-3 所示。

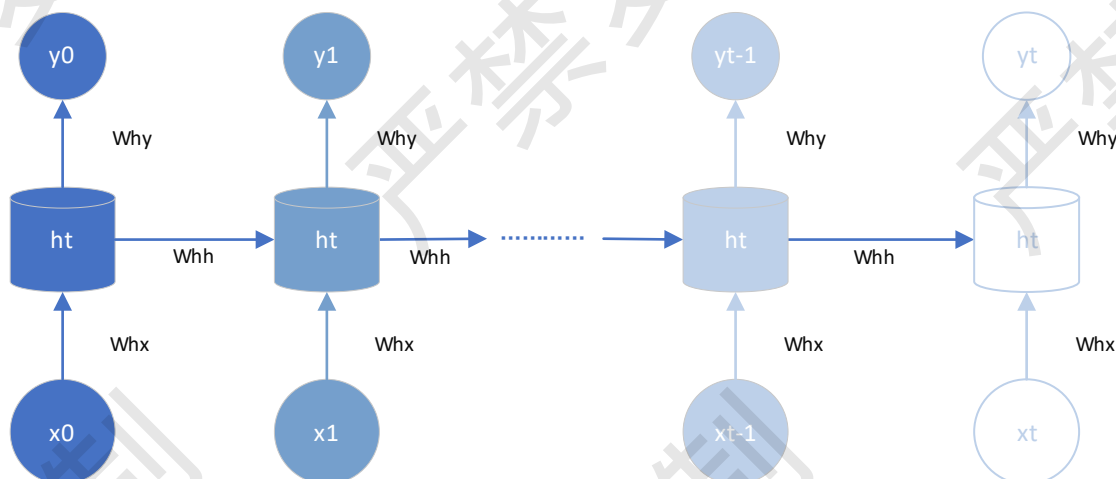


图 2-3 RNN 神经网络的梯度消失问题

图 2-3 中各个神经元的颜色深度表示与 0 时刻输入的感应程度。假如实际应用中 Y1 和 y0 时刻的感应程度较高，当使用循环神经网络结构进行训练师，Y1 对于 0 时刻的输入感应程度减弱，从而造成梯度消失的问题，导致网络性能不佳。

因此 RNN 神经网络对于长时间的山洪预警可能会出现梯度消失，可能会使输出的水位与真实的水位贴合程度较差，LSTM 神经网络的提出有效的改善了 RNN 神经网络所面临的梯度消失的问题。

2.3 LSTM 模型山洪预警分析

LSTM 是一种特殊的 RNN 神经网络，可以很好的解决长时依赖，从而梯度消失的问题。LSTM 是 RNN 的一种变体，更高级的 RNN，同样可以有效的处理序列数据。简单来讲 LSTM 比 RNN 增加了对信息选择性存储的能力，LSTM 拥有门控装置，可以对信息进行挑选。GRU 和 LSTM 同为 RNN 的变体，LSTM 拥有两个门限，GRU 对 LSTM 进行简化，只拥有一个门限，但两者非常相近，国内外众多学者采用 LSTM 进行研究，对于 LSTM 认可度较高，因此本文选择 LSTM 作为研究模型。

2.3.1 LSTM 神经网络的原理

LSTM 通过它的“门控装置”有效的缓解了梯度消失的问题。LSTM 神经网络结构是基于存储器结构的递归记忆型神经网络中的一种，它可将当前节点的隐藏状态记录下来，LSTM 拥有独特的存储器单元机制，使得梯度消失和爆炸问题得到较好解决。LSTM 结构如图 2-4 所示。

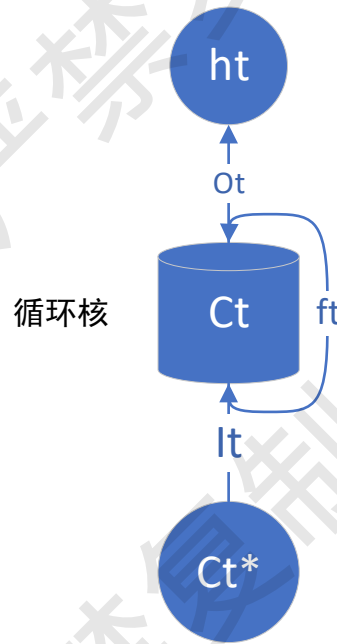


图 2-4 LSTM 的基本结构

LSTM 通过遗忘门 f_t 来决定我们从单元状态中舍弃哪些信息。输入 h_{t-1} 和 x_t 通过 sigmoid 激活函数，输出 0-1 之间的数字，来决定该遗忘哪些信息，如公式 2.3 所示。

$$f_t = \text{sigmoid}(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2.3)$$

LSTM 通过输入门决定我们应记录哪些信息，更新状态单元中的哪些信息。公式如 2.4, 2.5。

$$I_t = \text{sigmoid}(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2.4)$$

$$C_t^* = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (2.5)$$

不断更新单元状态，将旧的状态 C_{t-1} 更新为新的状态 C_t 。公式如 2.6。

$$C_t = f_t \cdot C_{t-1} + (1 - f_t) \cdot C_t^* \quad (2.6)$$

最后，决定要输出的内容，根据单元状态，输出经过过滤的版本， b_i, b_f 和 b_o 是待训练偏置项。我们将运行一个 sigmoid 激活函数，决定要输出单元哪一部分内容。然后将单元状态经过 tanh 激活函数，并将结果乘以 sigmoid 的输出，这将只输出我们决定的部分。公式如 2.7, 2.8。

$$O_t = \text{sigmoid}(W_o[h_{t-1}, x_t] + b_o) \quad (2.7)$$

$$h_t = O_t * \tanh(C_t) \quad (2.8)$$

2.3.2 LSTM 模型山洪预警的优势

LSTM 相比于上一节所介绍的 RNN 神经网络增加了门控装置，对信息进行挑选，并且使用累加的形式计算状态，这种累加形式使导数也是累加形式存在，从而避免了传统 RNN 神经网络因为链式求导法则而直接导致的梯度被连续乘积，导致梯度消失的问题。简单来说就是避免了 RNN 神经网络因为许多个小于 1 的项相乘导致梯度无限接近于 0 造成梯度消失的问题。当输入的水位、雨量、空气湿度、温度数据量较少时，RNN 神经网络可能并不会因为梯度消失问题而出现很大的劣势，当数据量较大，并且对准确度要求更高时，LSTM 神经网络会因为避免梯度消失问题的发生从而产生出明显的优势，水位预测的准确性会比 RNN 更加高。

2.4 本章小结

本章主要介绍苕溪小流域的水文特征，以及选取站点附近的特征，对模型选择进行了分析，介绍了 RNN 神经网络和 LSTM 神经网络的区别，介绍了 RNN 神经网络的缺点，并且由此诞生了 LSTM 神经网络，尽可能的解决了 RNN 神经网络所存在的问题。分析两者的原理，得出了 LSTM 神经网络基本在任何领域都优于传统的 RNN 神经网络。

第3章 小流域山洪预警模型设计

拟合自然数据是建模的最终目标，对于山洪预警技术的研究，往往采用水文数据模型或多种模型组合的计算方式。水文数据建模需要大量的基础资料，以及专业知识，并且规模较为庞大，涉及多个学科。而神经网络模型所需要数据量较小，可以根据获得数据送入神经网络自动拟合自然数据进行预测。神经网络根据过去，现在的数据对未来进行预测，可以大致的确定未来的发展趋势，方向等，甚至在短时间内可以进行较为准确的预测。相对于其他预测方式具有较强的优越性。应用与预测方面的神经网络主要有 BP 神经网络，RNN 神经网络和 LSTM 神经网络。本文主要提出的 RNN 神经网络和 LSTM 神经网络建立的山洪预警模型，依据山洪暴发的原因分析出所应选取的输入参数。通过时间预测来输出水位高度，并且对各个模型连续 48 小时预测性能进行对比，选出最优秀的模型进行水位的预测。

3.1 模型输入输出参数确定

根据山洪灾害发生的情况，结合临安地区的地形特征以及植被情况，连续暴雨，水位上升等情况，判断出山洪暴发的影响因子。影响因子主要有水位、雨量、空气湿度、温度、土壤含水量等因素。根据水位、雨量、空气湿度、温度来建立山洪预警模型。

核心算法主要是时间序列预警模型，将水位、雨量、空气湿度、温度，作为输入参数，将水位作为输出参数，建立以时间序列预测的山洪预警模型。根据预测的水位判断山洪预警的等级，进而更好的防范有可能发生的山洪灾害。

3.1.1 预警指标输入参数确定

本文研究区域是苕溪的桥东村站，桥东村位于浙江省天目山区附近，距离南溪与南苕溪交汇点下游 500 米处，是浙江省山洪高风险地区。根据对山洪灾害的原因分析，引发山洪暴发的因子包括雨量、空气湿度、温度、土壤含水量、水位、地质灾害等。

持续性的降雨是引发山洪灾害的主要因素，连续性的降雨导致土壤含水量过大，土壤无法吸收过多的水分，导致水位上涨。连续性的暴雨导致空气湿度大，水分蒸发慢，水位土壤含水量以及水分蒸发较慢。水位上升是导致山洪暴发的直接因素，

水位过高，超过警戒水位，甚至超过保证水位，从而发生决堤，溢出等现象引发山洪暴发。

分析以上信息后本人从浙江省雨水情网站上获取了站点名为桥东村的水位以及雨量的信息，保障实验数据样本的真实性。如图 3-1 为 2020 年 6 月 1 日 0 点至 2020 年 6 月 5 日 0 点的水位信息。图 3-2 为 2020 年 6 月 1 日 0 点至 2020 年 6 月 5 日 0 点的雨量信息。

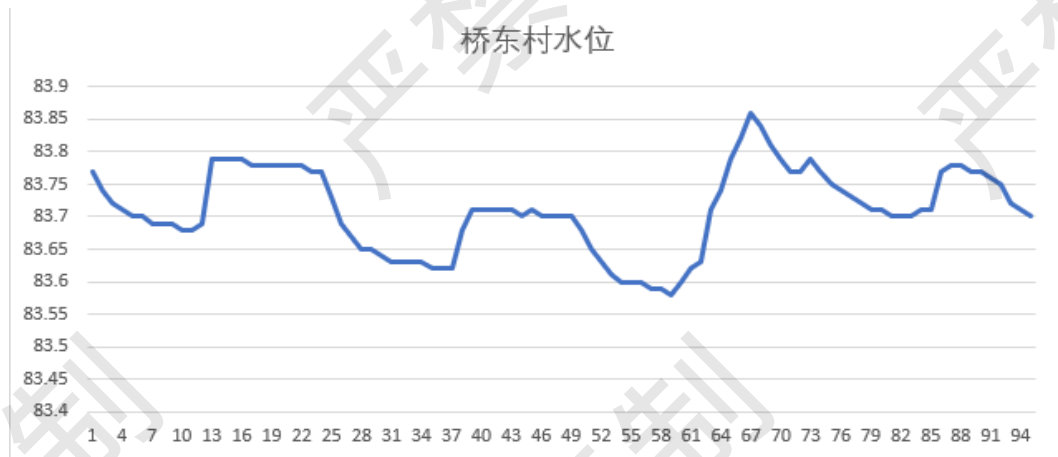


图 3-1 部分水位信息图

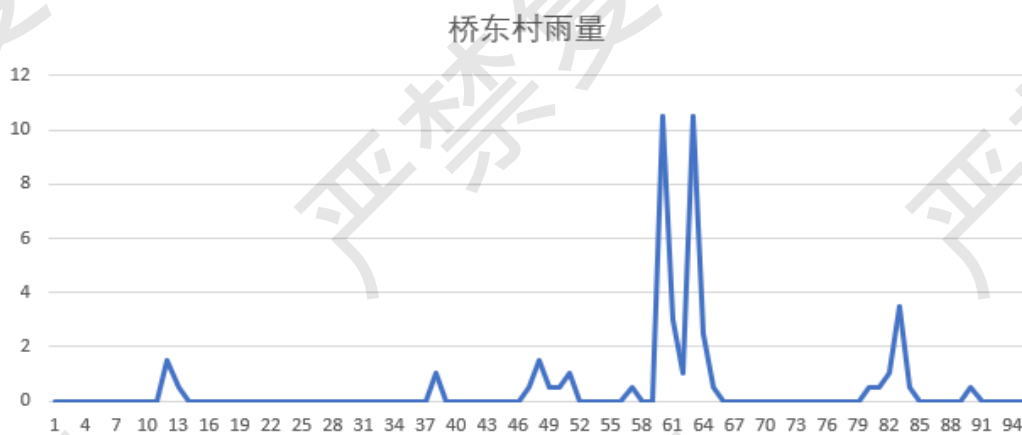


图 3-2 部分雨量信息图

从图中可以看出，当有雨量时，水位也会上涨。当发生连续性雨量时极有可能发生山洪灾害，因此水位和雨量时山洪暴发的主要因素。

当连续性湿润天气，且温度较低时，土壤中的水分蒸发量较少，当再次发生雨量时，土壤水分依旧处于饱和状态，从而导致土质较为疏松，进而可能发生山洪等灾害，因此空气含水量，气温可作为山洪暴发的影响因子之一。

3.1.2 预警信息输出参数确定

输出的参数即预警信息包括水位，雨量，空气湿度，气温作为输入参数，建立时间序列预测模型得到输出结果，即水位预测信息。在根据水位对应预警等级。预

警等级根据河道水位的警戒线以及保证水位线来评判和发布，将预警等级分为多个等分，并有防汛指挥部统一下发到各级部门以及人民群众中。

3.2 时间序列预测

基于 RNN 算法或 LSTM 的时间序列预测，即通过历史水位，雨量，空气湿度，温度去预测雨量和水位，实验设计通过对 2020 年 6 月 1 日~2021 年 4 月 1 日总共 7296 小时的数据进行模型训练，其中 5400 小时作为训练集，1896 小时作为测试集。通过连续输入 240 小时的数据来预测 241 小时的水位，以达到预测目的。

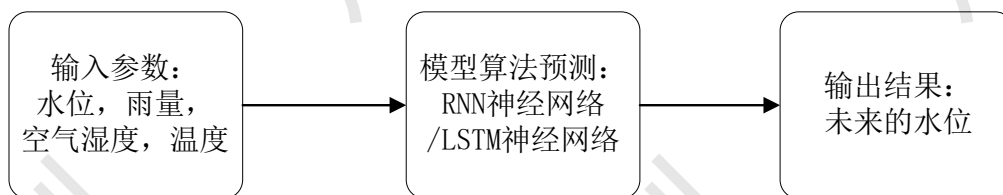


图 3-3 时间序列预测算法模型

3.3 模型实现流程

模型算法运行环境是 python3.7 版本，利用 tensorflow2.0 库在 jupyter notebook 平台实现 RNN 神经网络以及 LSTM 神经网络。

神经网络模型在 python 中的实现过程流程图及其步骤如图 3-4 所示。

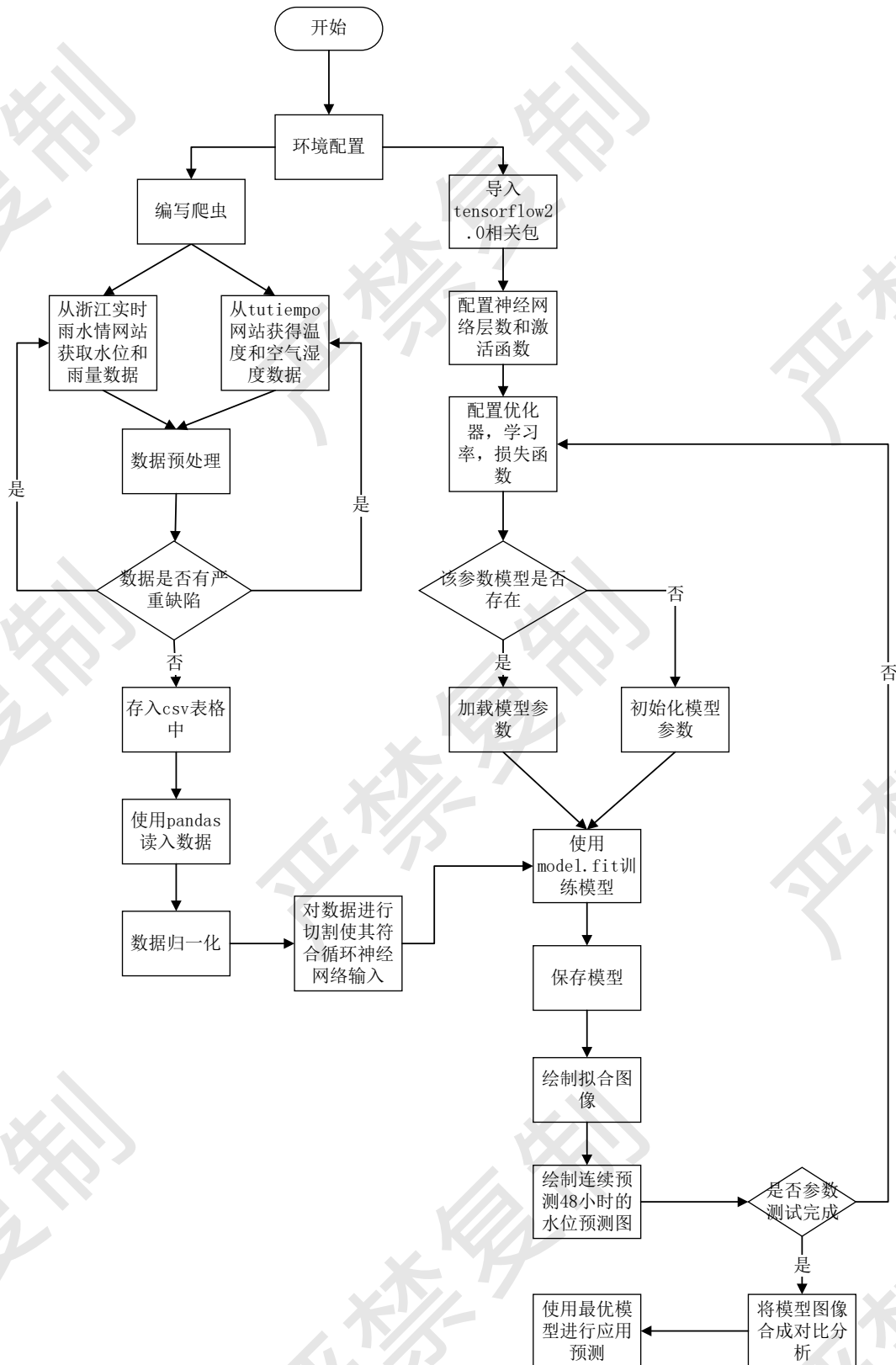


图 3-4 模型流程图

(1) 使用 anaconda 创建一个 python 版本为 3.7, 配置有 tensorflow2.0 的环

境，防治对其他的环境造成污染。

- (2) 编写网络爬虫代码分别浙江实时雨水情网站和 tutiempo 网站获取水位、降水、温度、空气湿度等数据。对数据进行预处理，将 null 值的数据设置为周围 24 小时的平均值，存入 csv 表格中。
- (3) 导入 tensorflow2.0, matplotlib, numpy 等相关包。
- (4) 配置神经网络的层数，激活函数，学习率，损失函数，优化器。如果模型存在，则导入存在的模型，如果模型不存在，则对模型进行初始化。
- (5) 使用 pandas 从 csv 格式中读入相关数据，对数据进行归一化。使用配置好的神经网络模型对数据进行训练，输出水位数据。
- (6) 使用 matplotlib 分别绘制验证集与测试集预测水位的拟合曲线。
- (7) 使用双层 for 循环实现连续 48 小时的水位预测，使用 matplotlib 绘制预测水位与真实水位的对比情况。
- (8) 对比 RNN 与 LSTM 在最优参数下的模型，将其绘制在同一张图内，并且将误差值制作成表格。
- (9) 选取最优模型对最近的水位进行预测，并且绘制图像。

3.4 本章总结

本章主要阐述了基于神经网络的山洪预警模型的输入输出指标的选择。分析选择区域应选择的输入指标，并且确定将雨量，水位，空气湿度，气温作为输入指标，建立算法模型，得到预测水位结果，根据水位情况判断预警等级。

主要介绍了模型从数据获取、数据的预处理、数据存储、数据读入、数据归一化、模型配置、模型训练、模型测试以及模型对于短期内水位预测的实现过程进行了描述，对模型的整体框架进行了介绍了。

第4章 小流域模型的实现

本章内容主要根据上一章对 RNN 神经网络以及 LSTM 神经网络讨论结果，以临安市桥东村为研究区域，实现对以雨量，水位，空气湿度，温度作为输入参数，水位作为输出参数的时间序列预测。

4.1 模型的实现环境

实验环境为 window10，64 位操作系统，处理器为 AMD Ryzen 7 4800U with Radeon Graphics，内存 16GB。python 为 3.7 版本，使用 jupyter notebook 平台作为代码编写平台，RNN、LSTM 神经网络均使用 tensorflow2.0 框架进行编写。

4.2 数据的获取

实验数据使用 python 的 requests 库，以及 xpath 库编写代码获取，水位以及雨量信息从浙江实时雨水情预警平台获得，湿度以及气温的历史信息从 tutiempo 网站获取。

4.2.1 水位和雨量信息获取

根据上一章分析可知，雨量和水位对于山洪有较为密切的影响，获取水位以及雨量数据需要如下步骤：（1）选择临安市附近较为合适的水位站作为具体的研究对象。（2）编写网络爬虫代码，将获取来的数据放入 txt 文件中。（3）编写代码对获取来的数据进行提取以及清洗，并且将数据放入 csv 文件中。

根据分析选择临安市的桥东村水位站作为研究对象较为合适，桥东村水位站两面环山，当发生持续性降水时容易从山地汇聚大量雨水，从而导致河道水位上升，进而超过水位线发生山洪。桥东村水位站的警戒水位线是 84.66 米，保证水位线是 85.66，可以根据警戒水位线以及保证水线来判断应发布的山洪预警等级。

桥东村水位站的 id 为 63000200，使用网络爬虫爬取该站点从 2020 年 6 月 1 日至，2021 年 4 月 1 日的数据，保存到 63000200 桥东村.txt 里面。编写 python 代码将爬取到的数据转化为字典格式，并且将水位和雨量数据为 null 的初始化为 0，将数据保存到 63000200 桥东村.csv 里面。打开 63000200 桥东村.csv，通过筛选观察

被赋值为 0 的数据，将其修改为最近 24 小时的平均值，从而使数据更加具有准确性。

4.2.2 空气湿度温度信息获取

空气湿度和温度会影响水分的蒸发量，从而影响土壤中的含水量以及在晴朗天气水位的下降速度，从而空气湿度和温度也会对山洪暴发产生间接性的影响。空气湿度和温度的数据均来自 tutiempo 网站。该网站记录了全世界各大城市多年的温度，天气，风力，空气湿度以及风向等信息。选择中国浙江省杭州市，将空气湿度以及温度的数据从 2020 年 6 月 1 日至 2021 年 4 月 1 日的数据使用网络爬虫代码填入 63000200 桥东村.csv 的空气湿度以及温度栏中。为了保证数据的准确性，进行人工检查，将一些网站未录入数据的空气湿度以及温度修改为最近 7 天的空气湿度以及温度，保证数据尽可能的可靠。

4.3 模型的代码的编写

4.3.1 数据读入以及归一化

数据使用 python 的 pandas 库进行读入，将数据保存到 sw 数组中，使用 iwoc 函数将 sw 数组分割为 training_set 和 test_set。training_set 保存训练集数据，test_set 保存测试集数据。将数据集与测试集均分割为 x_train_set 以及 y_train_set 和 x_test_set 以及 y_test_set。使用 fit_transfrom 和 transform 分别对 x 和 y 进行归一化。分开归一化有助于后续进行反归一化还原数据。

4.3.2 构造神经网络模型

使用 for 循环将 5400 个训练数据按照连续 240 小时一组分割为 5160 组，数据维度为(5160,240,4)，将 1896 个测试数据按照连续 240 小时一组分割为 1856 组，数据维度为(1856,240,4)。RNN 神经网络的输入要求是[送入样本数，循环核时间展开步数，每个时间步输入特征个数]，数据维度符合要求。使用一层带有 100 个循环核的 RNN 神经网络，选择 tanh 激活函数，选择 Adam 优化器，损失函数使用均方误差，设置保存最佳权重，每次送入机器的 batch_size 设置为 64。使用 fit 函数对模型进行拟合。

4.3.3 绘制预测图像

使用 matplotlib 库绘制图像，将训练集损失大小和测试集损失大小通过调用 history 从模型中提取出来，再使用 plot 以及 show 函数绘制误差值的大小。

使用模型预测训练集的数据，将预测的水位和真实水位都进行反归一化。使用 plot 函数和 show 函数绘制训练集的拟合图像以及测试集的拟合图像。

4.3.4 连续预测 48 小时的水位

根据已经训练完的模型对 48 小时数据进行预测，首先使用最近的 240 小时预测第 241 小时的数据，再将预测的第 241 小时的数据加上后 239 小时预测第 242 小时的数据，再使用后 238 小时加上预测的 241 和 242 小时预测第 243 小时的数据，以此类推共预测 48 小时。将预测的 48 小时和真实数据的 48 小时使用 plot 函数和 show 函数绘制图像进行对比，并且计算各种误差。

4.4 水位的时间序列预测

实验将通过 RNN 神经网络算法，LSTM 神经网络算法进行时间序列预测。实验步骤如下：（1）对数据进行归一化，将数据集分割为训练集和测试集。（2）将训练集和测试集连续 240 小时的数据作为一个组，第 241 小时作为该组的标签，循环核时间展开步为 240。（3）将训练集输入神经网络进行训练，保存训练参数，保存最优权重。（4）根据训练好的模型，将测试集数据输入模型进行预测。根据预测结果绘制预测水位与真实水位的拟合图像。输出均方误差，均方根误差，平均绝对误差。（5）通过对比真实水位与拟合水位的情况，检验算法的优劣，不断修改模型的参数，尽可能的寻找最优解。对比 RNN 神经网络与 LSTM 神经网络的优劣。

4.4.1 RNN 神经网络的水位预测

训练参数设置如表 4-1 所示。

表 4-1 训练参数

参数/取值	参数[1]	参数[2]	参数[3]	参数[4]
学习率	0.05	0.01	0.005	0.001
Epoch	50	150	250	400

当 RNN 神经网络选取的训练参数[1]为:学习率设置为 0.05，训练进行 50 个 epochs，实验步骤如 4.3 章节介绍，训练数据为 5400 个小时的数据作为训练样本进行 RNN 神经网络模型训练，训练完毕对后 1896 小时的数据进行预测，并对预测数据与真实数据进行对比。图 4-1 为训练集水位预测模型结果，图 4-2 为测试集水位预测模型结果，图 4-3 为训练集损失函数大小。从图书可以看出在训练参数[1]的情况下，模型出现欠拟合的情况，对于测试集的拟合情况较差。

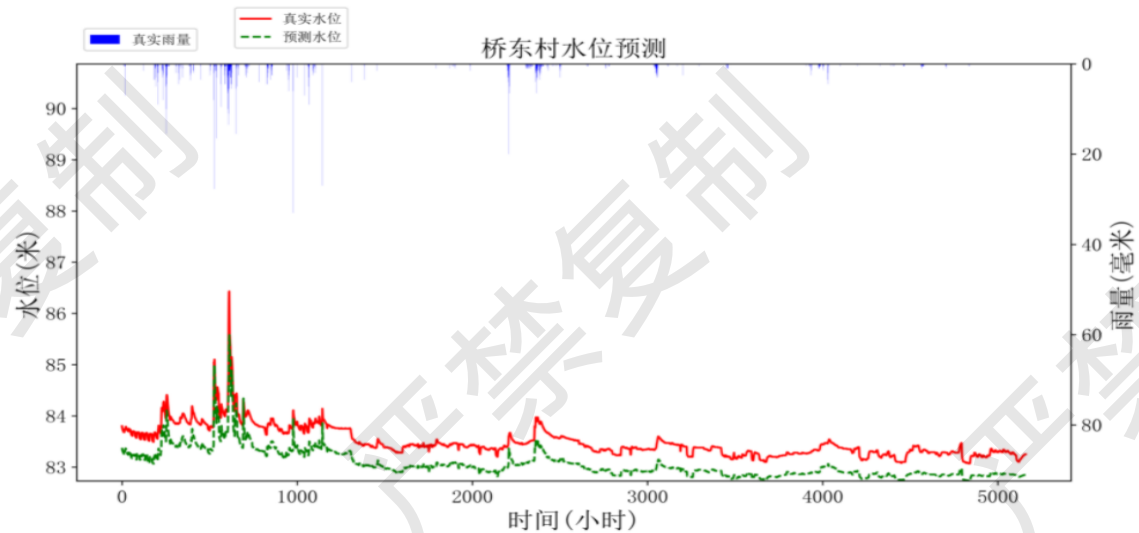


图 4-1 RNN 参数[1]训练集拟合

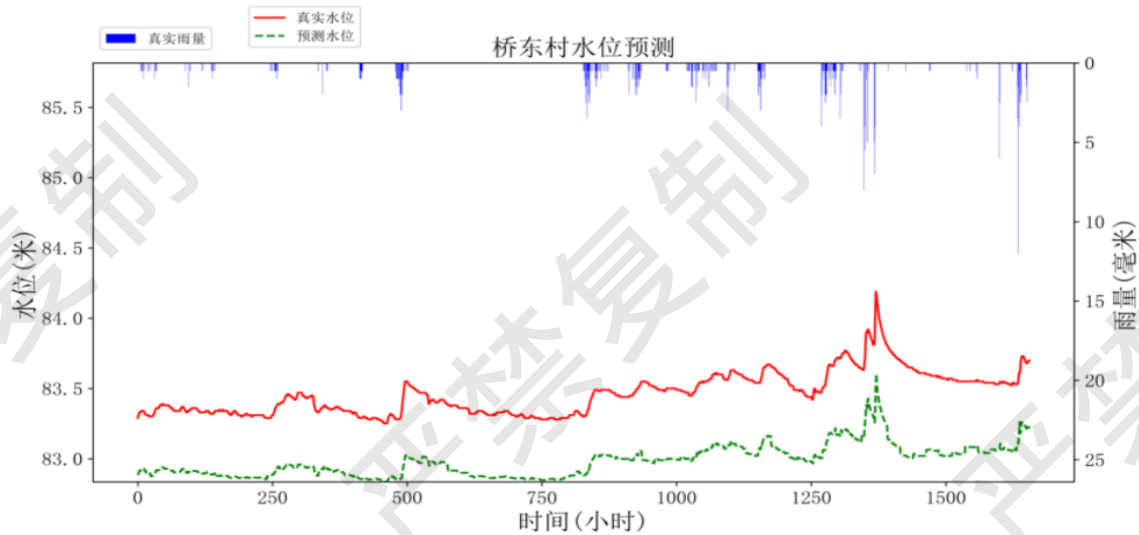


图 4-2 RNN 参数[1]测试集拟合

均方误差: 0.232198

均方根误差: 0.481869

平均绝对误差: 0.479251

图 4-3 RNN 参数[1]测试集水位误差

RNN 神经网络选取的训练参数[2]为: 学习率设置为 0.01, 训练进行 150 个 epochs, 训练数据为 5400 个小时的数据作为训练样本, 后 1896 小时的数据进行预测。预测数据结果如图 4-4~4-6 所示。当选取训练参数为[2]时, 模型较参数[1]具有较高的提升, 但从图 4-6 可以看出对于测试集的拟合程度依旧较低。

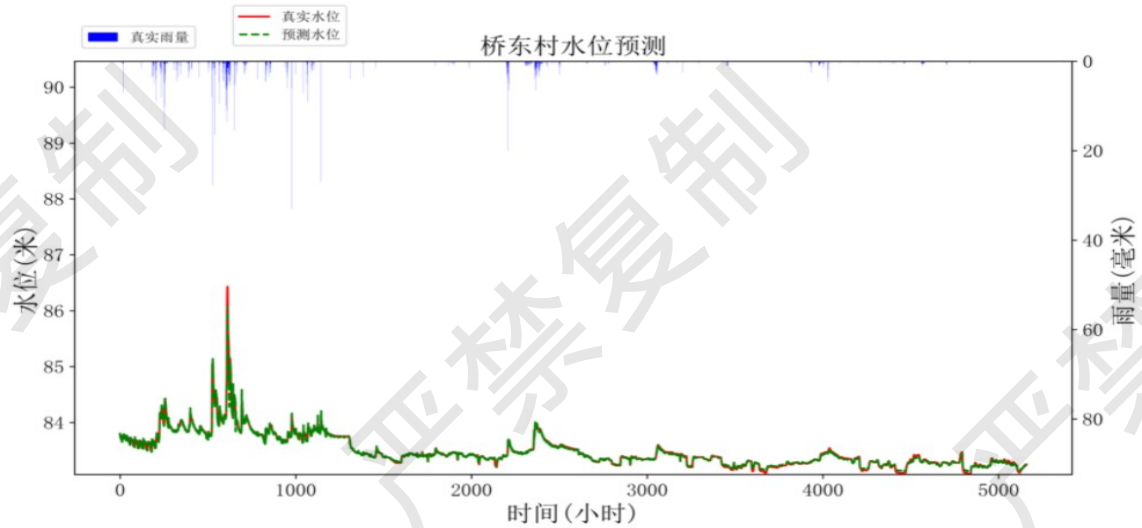


图 4-4 RNN 参数[2]训练集拟合

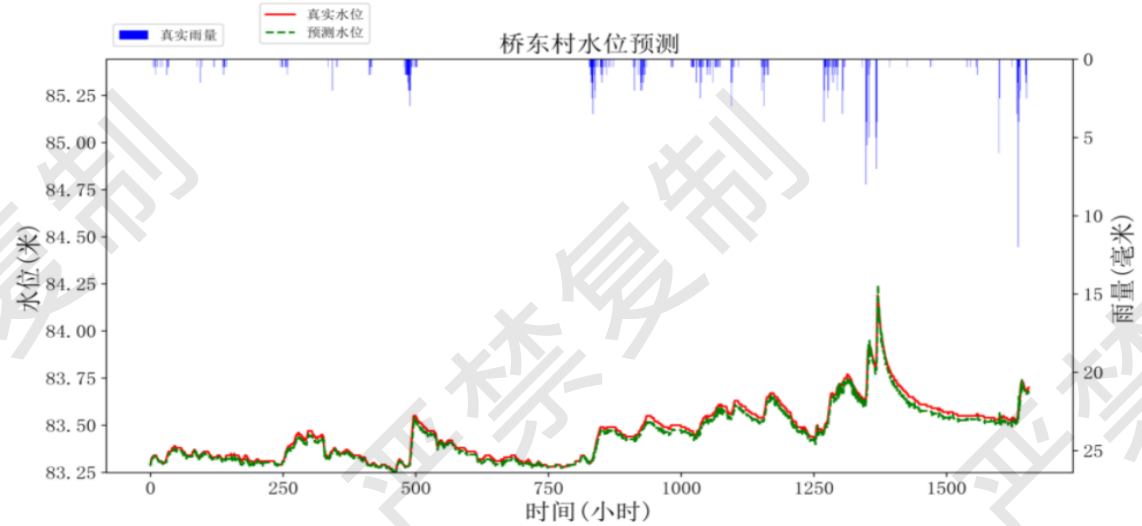


图 4-5 RNN 参数[2]测试集拟合

均方误差: 0.000493
 均方根误差: 0.022212
 平均绝对误差: 0.018693

图 4-6 RNN 参数[2]测试集水位误差

RNN 神经网络选取训练参数[3]为,学习率设置为 0.005,训练进行 250 个 epochs,训练数据为 5400 个小时的数据作为训练样本,后 1896 小时的数据进行预测。预测数据结果如图 4-7~4-9 所示。选取训练参数[3]进行训练,从图 4-8 中可以看出测试集的真实数据与预测数据拟合程度较高,训练取得较为优秀的成果。

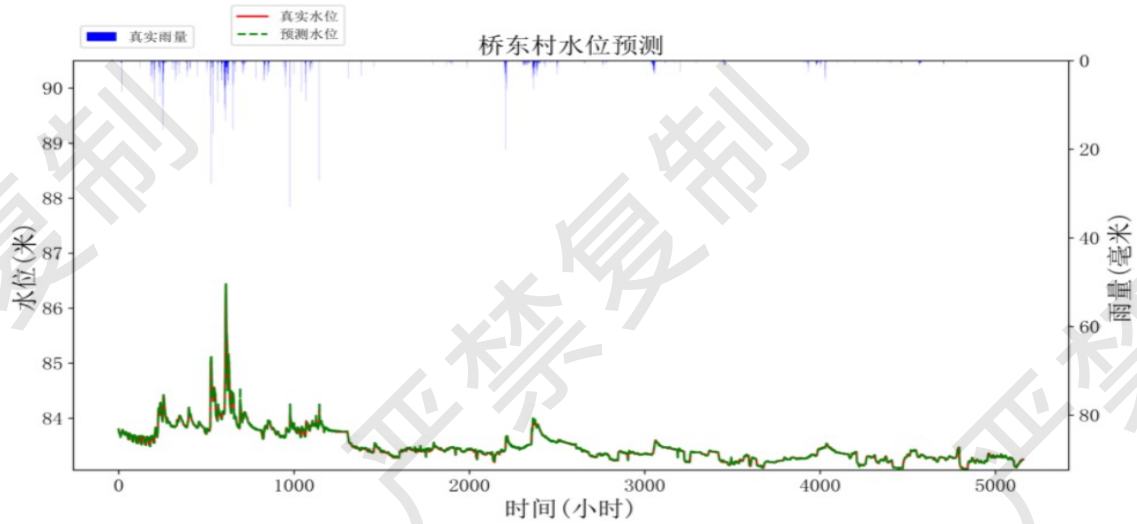


图 4-7 RNN 参数[3]训练集拟合

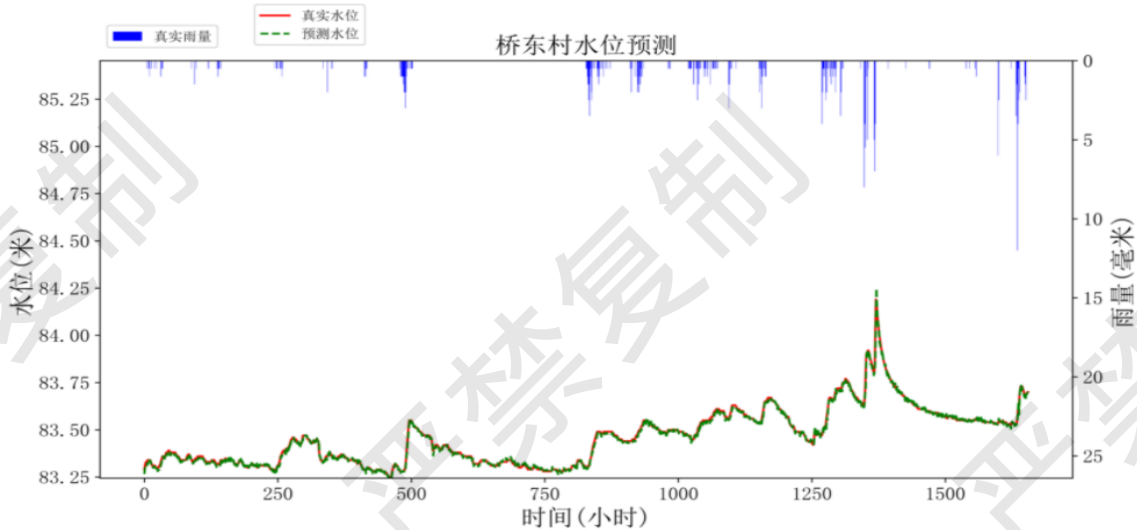


图 4-8 RNN 参数[3]测试集拟合

均方误差: 0.000136
 均方根误差: 0.011650
 平均绝对误差: 0.008529

图 4-9 RNN 参数[3]测试集水位误差

RNN 神经网络选取训练参数[4]为,学习率设置为 0.001,训练进行 400 个 epochs,训练数据为 5400 个小时的数据作为训练样本,后 1896 小时的数据进行预测。预测数据结果如图 4-10~4-12 所示。选取训练参数[4]进行训练,从图 4-10 和 4-11 可以看出模型取得较为优秀的拟合成果,误差值为 4 个模型中最低,均方误差为 0.000108,均方根误差为 0.010941,平均绝对误差为 0.007111。

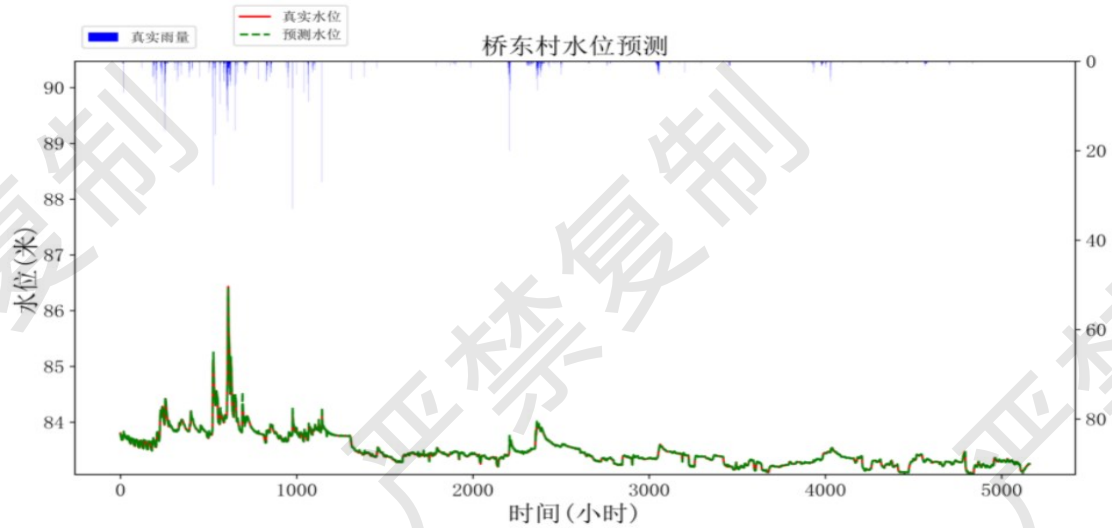


图 4-10 RNN 参数[4]训练集拟合

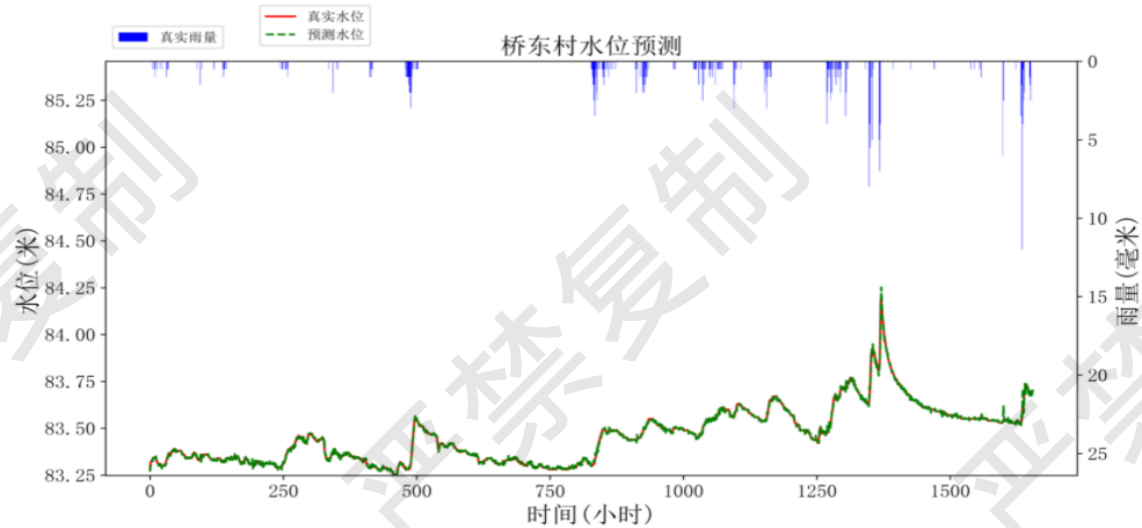


图 4-11

图 4-11 RNN 参数[4]测试集拟合

均方误差：0.000108

均方根误差：0.010393

平均绝对误差：0.007111

图 4-12 参数[4]测试集水位误差

表 4-2 RNN 神经网络水位模型误差对比

RNN 参数/水位 误差	RNN 参数[1] 学习率 0.05 步长 50	RNN 参数[2] 学习率 0.01 步长 150	RNN 参数[3] 学习率 0.005 步长 250	RNN 参数[4] 学习率 0.001 步长 400
均方误差	0.232198	0.000493	0.000136	0.000108
均方根误差	0.481869	0.022212	0.011650	0.010393
平均绝对误差	0.479251	0.018693	0.008529	0.007111

实现表明 RNN 神经网络在学习率较大，且步长较短的情况下，拟合程度较差，因学习率过大梯度下降法很难拟合成果。在学习率为 0.001，步长为 400 时，在训练

集以及测试集上误差均较小，取得实验中最好的效果。

4.4.2 LSTM 神经网络的水位预测

LSTM 神经网络选取训练参数[1]为，学习率=0.05，训练次数 50 次，训练数据为 5400 个小时的数据作为训练样本，后 1896 小时的数据进行预测。预测数据结果如图 4-13~4-15 所示。从图 4-14 以及图 4-15 可以看出 LSTM 神经网络在当前数据的情况下较 RNN 神经网络更加优秀。

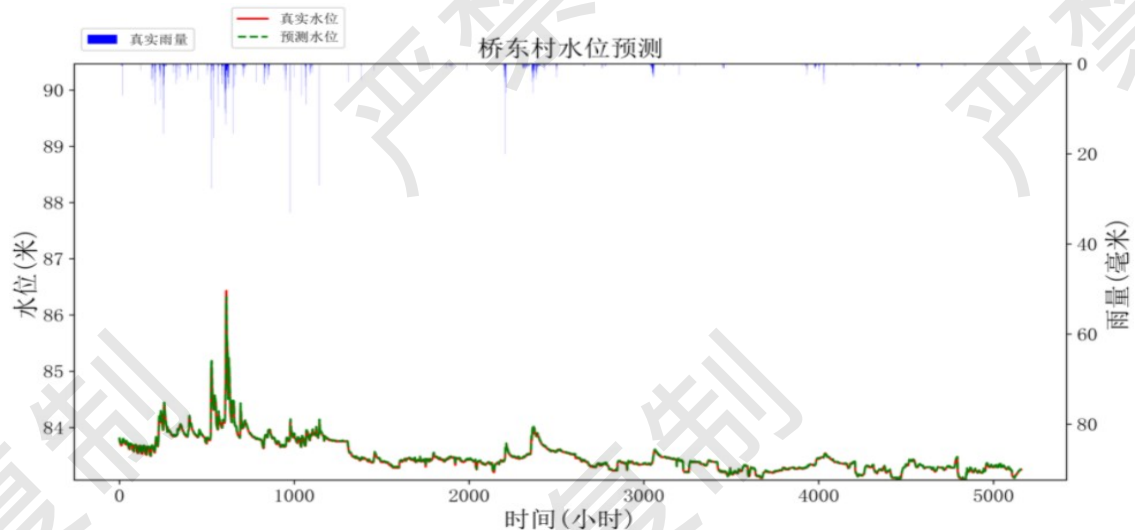


图 4-13 LSTM 参数[1]训练集拟合

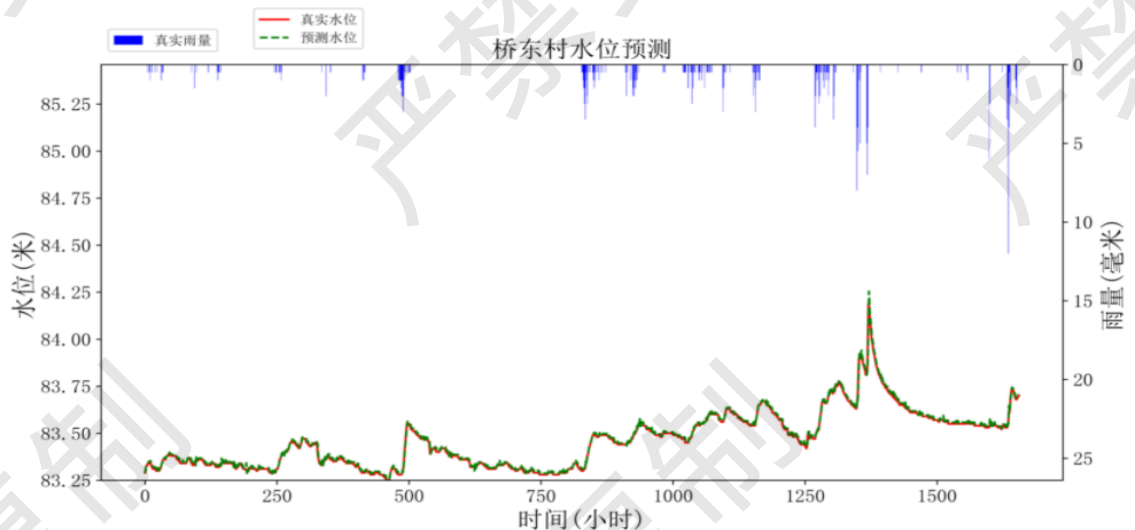


图 4-14 LSTM 参数[1]测试集拟合

均方误差：0.000138
均方根误差：0.011740
平均绝对误差：0.009235

图 4-15 参数[1]测试集水位误差

LSTM 神经网络选取训练参数[2]为，学习率=0.01，训练次数 150 次，训练数据

为 5400 个小时的数据作为训练样本，后 1896 小时的数据进行预测。预测数据结果如图 4-16~4-18 所示。

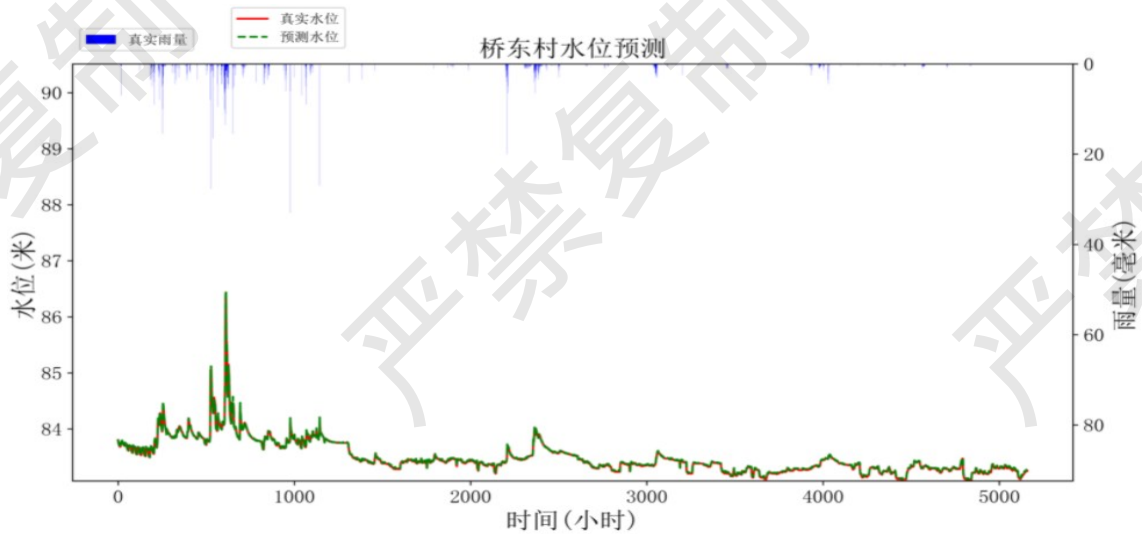


图 4-16 LSTM 参数[2]训练集拟合

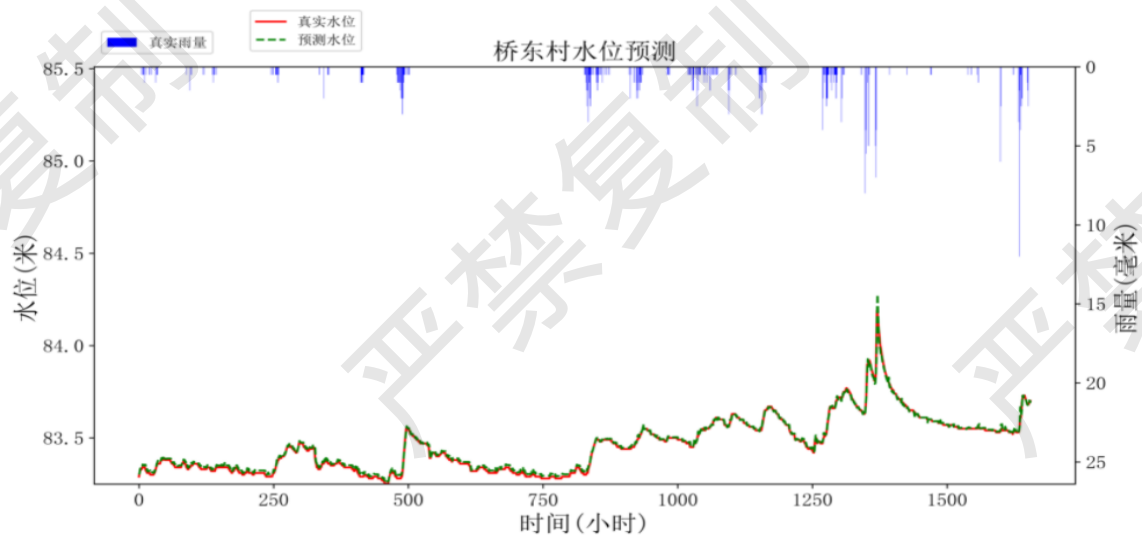


图 4-17 LSTM 参数[2]测试集拟合

均方误差: 0.000138
均方根误差: 0.011747
平均绝对误差: 0.009214

图 4-18 参数[2]测试集水位误差

LSTM 神经网络选取训练参数[3]为，学习率=0.005，训练次数为 250 次，训练数据为 5400 个小时的数据作为训练样本，后 1896 小时的数据进行预测。预测数据结果如图 4-19~4-21 所示。

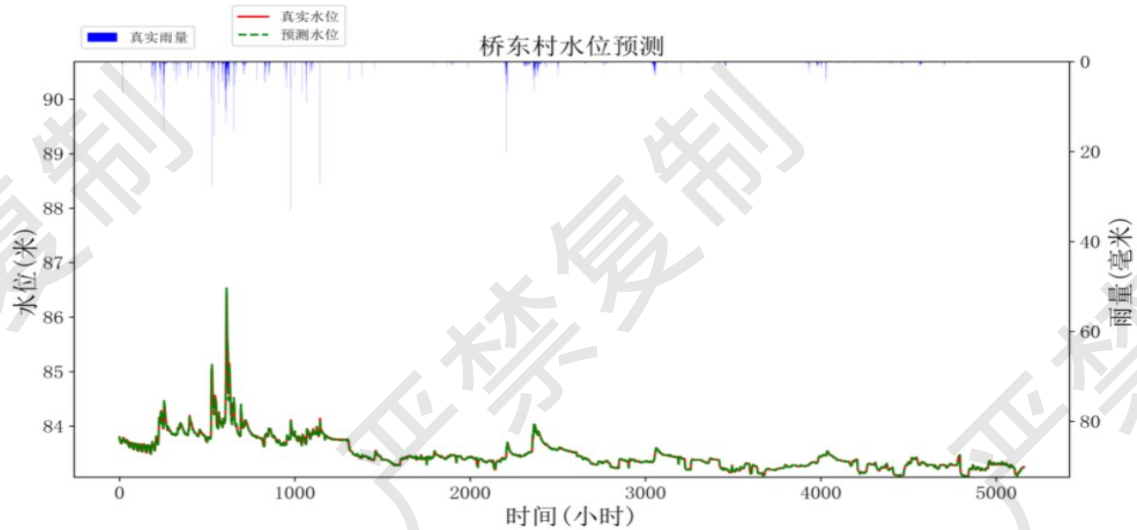


图 4-19 LSTM 参数[3]训练集拟合

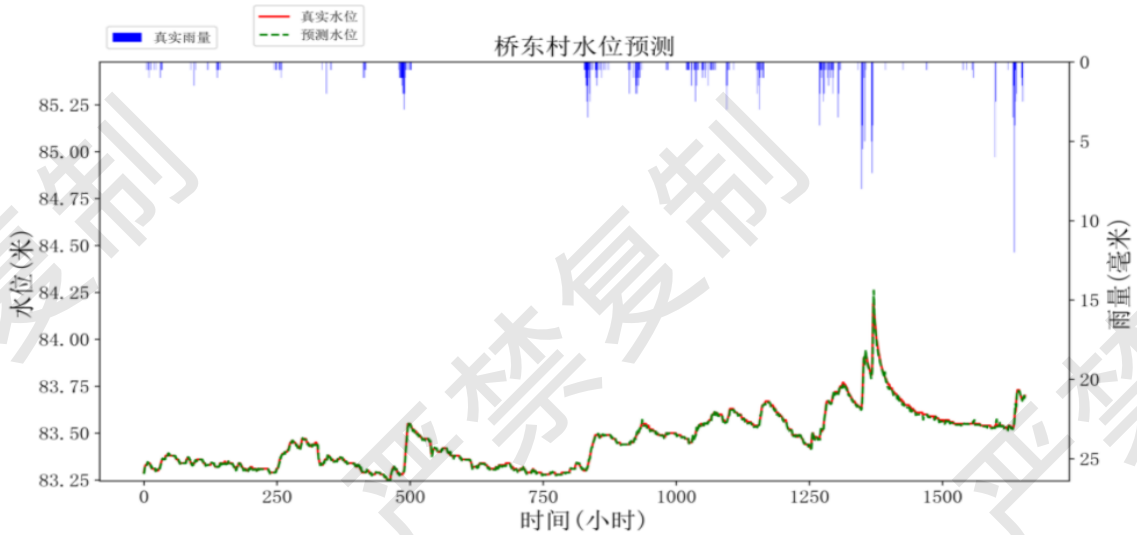


图 4-20 LSTM 参数[3]测试集拟合

均方误差: 0.000079

均方根误差: 0.008887

平均绝对误差: 0.006173

图 4-21 参数[3]测试集水位误差

LSTM 神经网络选取训练参数[4]为, 学习率=0.001, 训练次数为 400 次, 训练数据为 5400 个小时的数据作为训练样本, 后 1896 小时的数据进行预测。预测数据结果如图 4-22~4-24 所示。

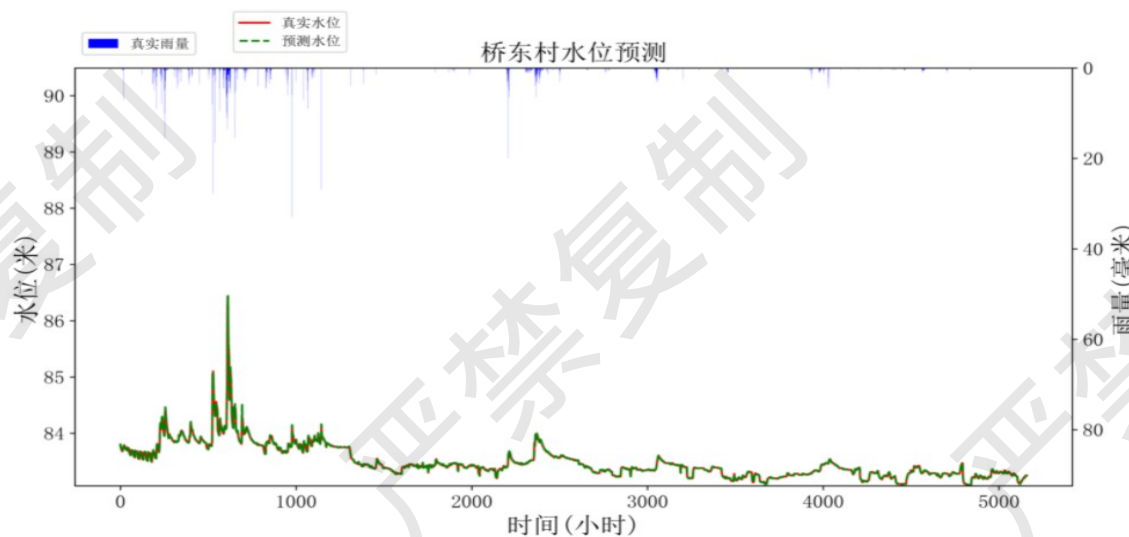


图 4-22 LSTM 参数[4]训练集拟合

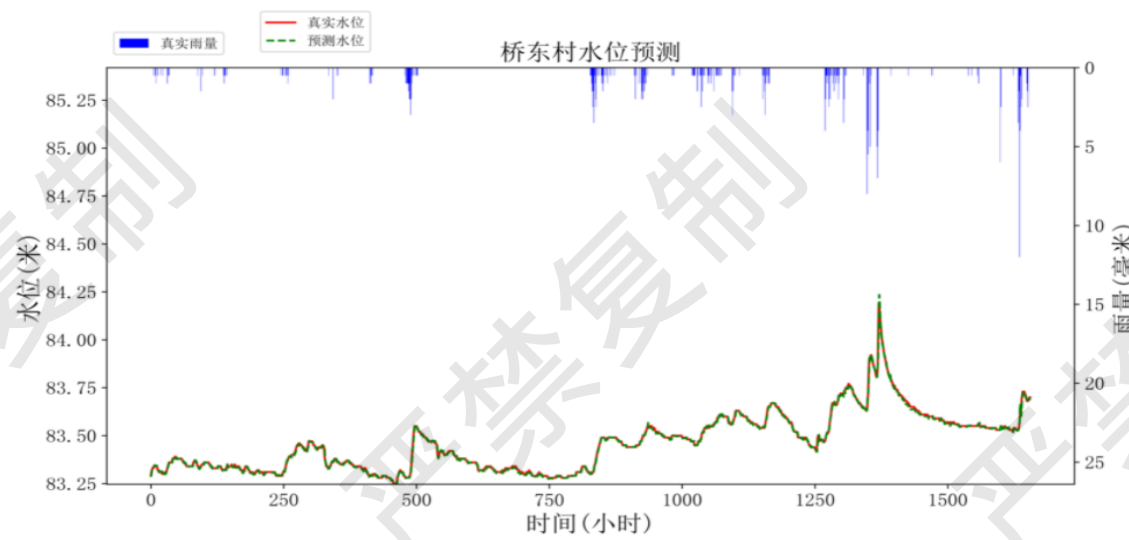


图 4-23 LSTM 参数[4]测试集拟合

均方误差：0.000071
均方根误差：0.008442
平均绝对误差：0.005551

图 4-24 参数[4]测试集水位误差

表 4-3 LSTM 神经网络水位模型误差对比

LSTM 参数/水位误差	LSTM 参数[1] 学习率 0.05 步长 50	LSTM 参数[2] 学习率 0.01 步长 150	LSTM 参数[3] 学习率 0.005 步长 250	LSTM 参数[4] 学习率 0.001 步长 400
均方误差	0.008471	0.000293	0.000079	0.000071
均方根误差	0.092039	0.017117	0.008887	0.008442
平均绝对误差	0.074939	0.015316	0.006531	0.005551

实验表明 LSTM 在长时间的水位预测上效果均较好，在学习率为 0.001，步长为 400 时拟合效果最好，模型最佳。LSTM 模型在长时间的水位预测的准确率上效果好于 RNN 神经网络。

4.5 模型预测近期水位

实验将通过 RNN 神经网络与 LSTM 神经网络进行连续 48 小时的水位预测，从而为山洪防治提供更加充裕的时间。实验的步骤如下例所述：（1）通过对 5400 小时的数据进行训练，分别得到 RNN 神经网络与 LSTM 神经网络各种参数的模型。（2）将本文选取的测试集数据中的 2021 年 3 月 19 日至 3 月 29 日共计 240 小时的各项数据作为输入参数送入模型进行预测得到第 241 小时的水位预测数据。再将预测的第 241 小时的数据加上后 239 小时预测第 242 小时的数据，再使用后 238 小时加上预测的 241 和 242 小时预测第 243 小时的数据，以此类推共预测 48 小时。（3）将预测出来的 48 小时的水位数据与 3 月 30 日与 3 月 31 日真实的水位数据进行对比。

4.5.1 RNN 神经网络预测近期水位

当 RNN 神经网络选取的训练参数[1]为:学习率设置为 0.05，训练进行 50 个 epochs，实验步骤如 4.4 章节介绍。连续 48 小时水位预测数据结果如图 4-25~4-26 所示。

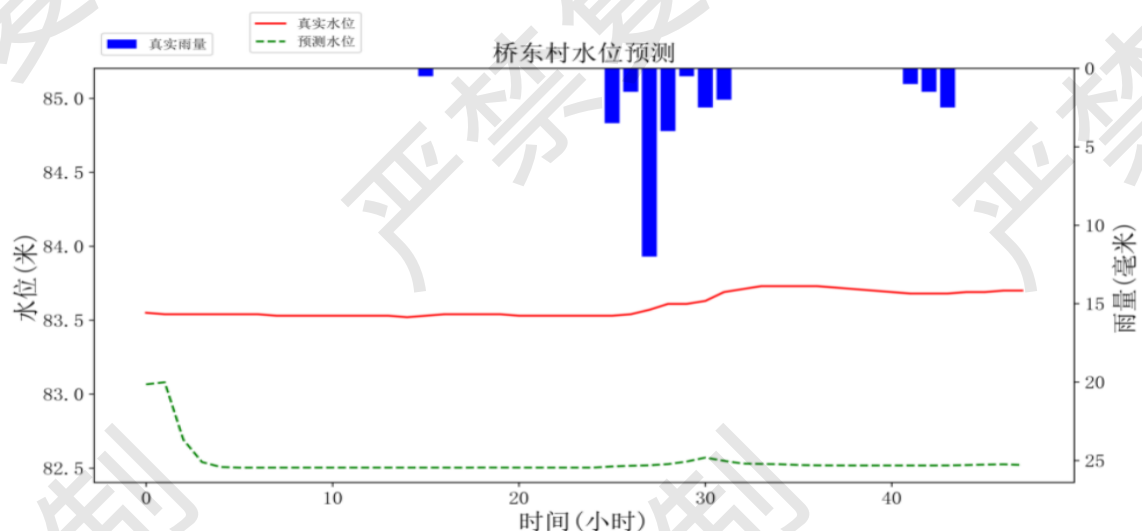


图 4-25 RNN 参数[1]48 小时水位预测

均方误差: 1.141168
均方根误差: 1.068255
平均绝对误差: 1.058277

图 4-26 RNN 参数[1]48 小时水位预测误差

当 RNN 神经网络选取的训练参数[2]为:学习率设置为 0.01，训练进行 150 个 epochs，实验步骤如 4.4 章节介绍。连续 48 小时水位预测数据结果如图 4-27~4-28 所示。

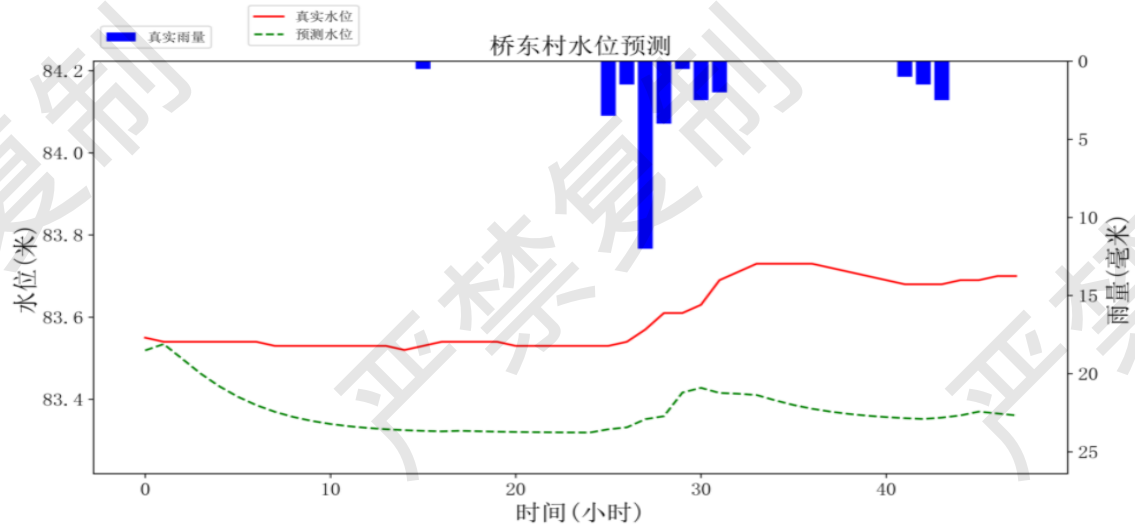


图 4-27 RNN 参数[2]48 小时水位预测

均方误差: 0.060751
 均方根误差: 0.246477
 平均绝对误差: 0.230145

图 4-28 RNN 参数[2]48 小时水位预测误差

当 RNN 神经网络选取的训练参数[3]为:学习率设置为 0.005, 训练进行 250 个 epochs, 实验步骤如 4.4 章节介绍。连续 48 小时水位预测数据结果如图 4-29~4-30 所示。

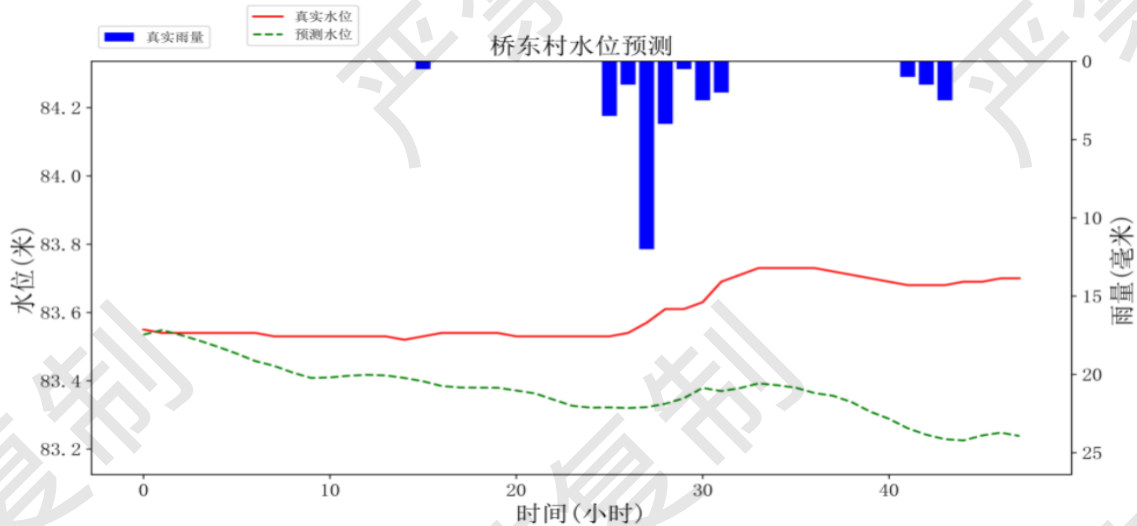


图 4-29 RNN 参数[3]48 小时水位预测

均方误差: 0.072093
 均方根误差: 0.268501
 平均绝对误差: 0.229147

图 4-30 RNN 参数[3]48 小时水位预测误差

当 RNN 神经网络选取的训练参数[4]为:学习率设置为 0.001, 训练进行 400 个 epochs, 实验步骤如 4.4 章节介绍。连续 48 小时水位预测数据结果如图 4-31~4-32 所示。

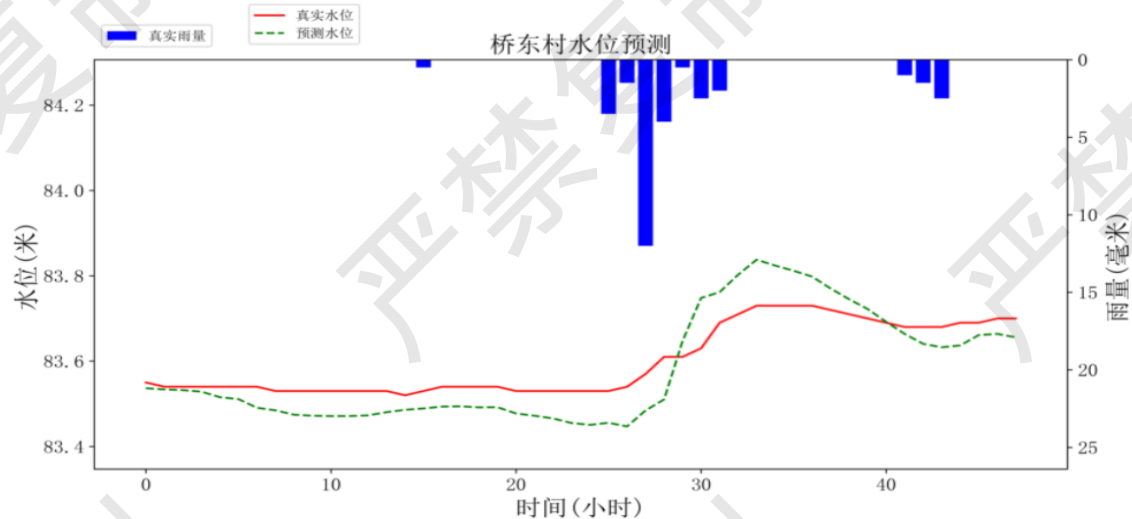


图 4-31 RNN 参数[4]48 小时水位预测

均方误差: 0.003473

均方根误差: 0.058928

平均绝对误差: 0.052403

图 4-32 RNN 参数[4]48 小时水位预测误差

表 4-4 RNN 神经网络连续预测误差对比

RNN 参数/水位 误差	RNN 参数[1] 学习率 0.05 训练次数 50 次	RNN 参数[2] 学习率 0.01 训练次数 150 次	RNN 参数[3] 学习率 0.005 训练次数 250 次	RNN 参数[4] 学习率 0.001 训练次数 400 次
均方误差	1.141168	0.107707	0.072099	0.003473
均方根误差	1.068255	0.317973	0.268501	0.058928
平均绝对误差	1.058277	0.302789	0.229147	0.052403

从图 4-25~图 4-32 以及表 4-3 可知, 在当前数据的情况下, 学习率为 0.001, 训练次数 400 次的情况下, 模型对于短期水位预测效果较为优秀。

4.5.2 LSTM 神经网络预测近期水位

当 LSTM 神经网络选取的训练参数[1]为:学习率设置为 0.05, 训练进行 50 个 epochs, 实验步骤如 4.4 章节介绍。连续 48 小时水位预测数据结果如图 4-33~4-34 所示。

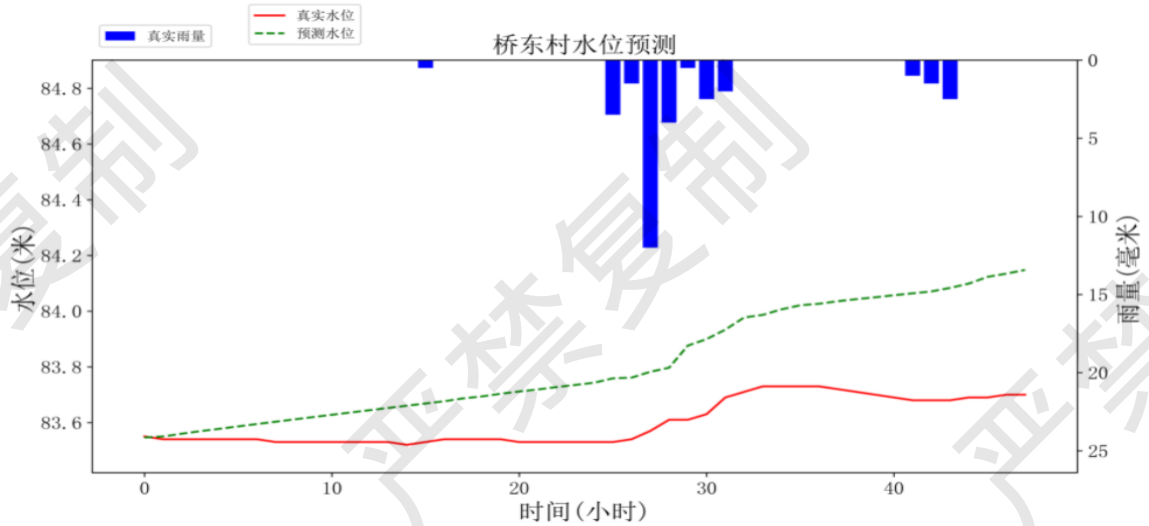


图 4-33 LSTM 参数[1]48 小时水位预测

均方误差: 0.059312

均方根误差: 0.243540

平均绝对误差: 0.209090

图 4-34 LSTM 参数[1]48 小时水位预测误差

当 LSTM 神经网络选取的训练参数[2]为:学习率设置为 0.01, 训练进行 150 个 epochs, 实验步骤如 4.4 章节介绍。连续 48 小时水位预测数据结果如图 4-35~4-36 所示。

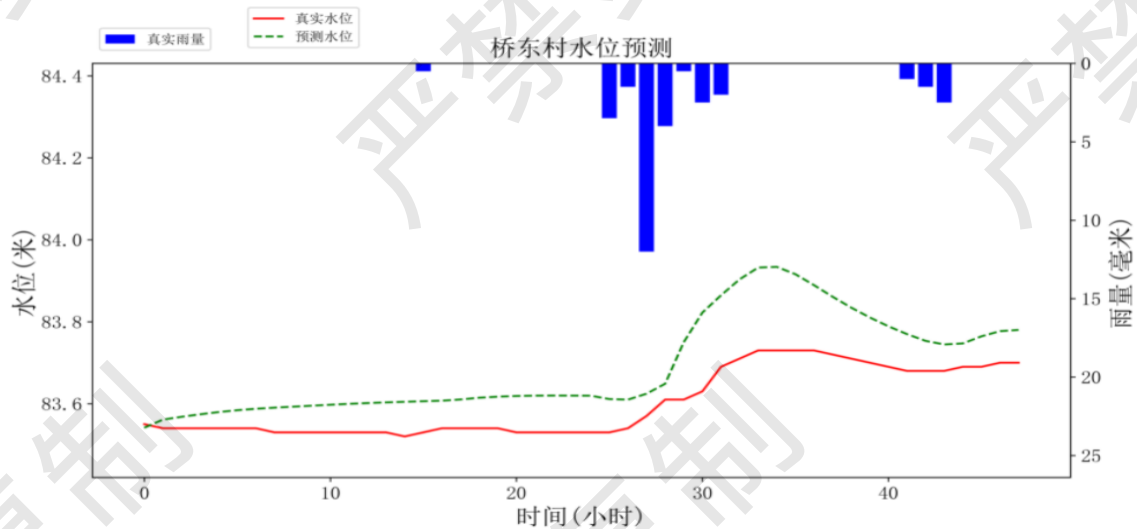


图 4-35 LSTM 参数[2]48 小时水位预测

均方误差: 0.010267

均方根误差: 0.101328

平均绝对误差: 0.089104

图 4-36 LSTM 参数[2]48 小时水位预测误差

当 LSTM 神经网络选取的训练参数[3]为:学习率设置为 0.005, 训练进行 250 个 epochs, 实验步骤如 4.4 章节介绍。连续 48 小时水位预测数据结果如图 4-37~4-38

所示。

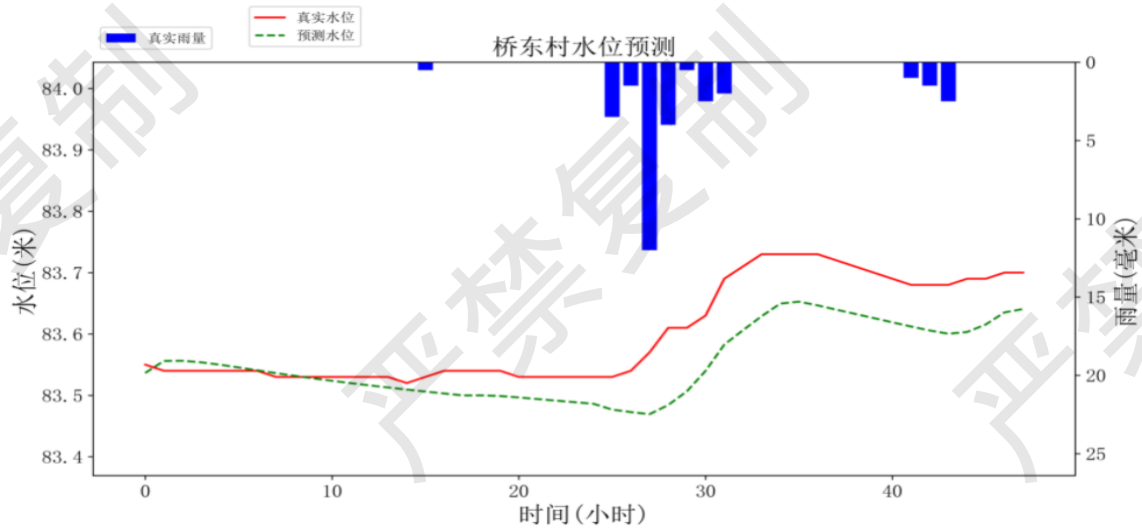


图 4-37 LSTM 参数[3]48 小时水位预测

均方误差: 0.003761

均方根误差: 0.061327

平均绝对误差: 0.050487

图 4-38 LSTM 参数[3]48 小时水位预测误差

当 LSTM 神经网络选取的训练参数[4]为:学习率设置为 0.001, 训练进行 400 个 epochs, 实验步骤如 4.4 章节介绍。连续 48 小时水位预测数据结果如图 4-39~4-40 所示。

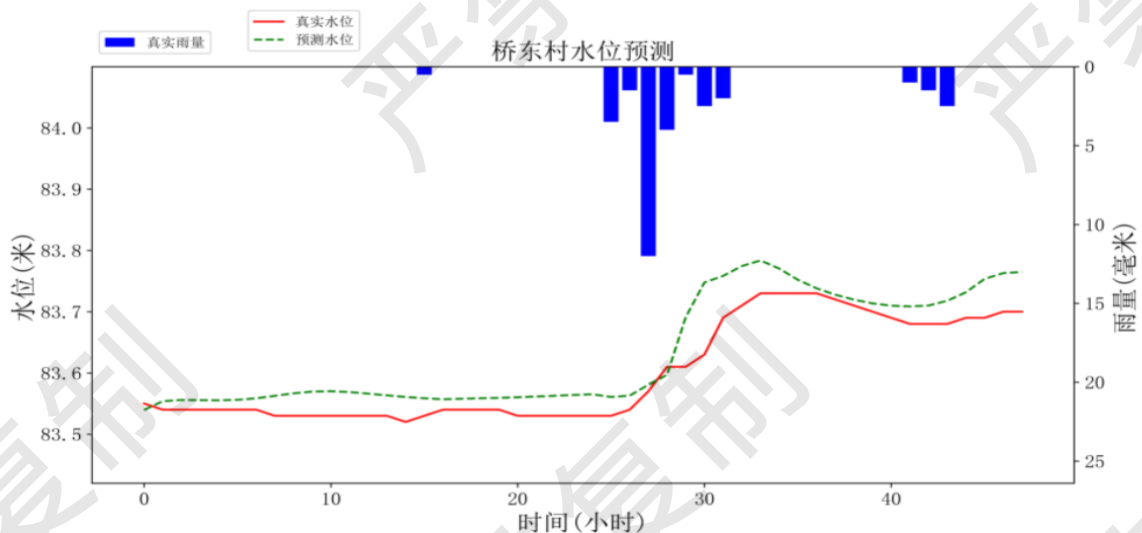


图 4-39 LSTM 参数[4]48 小时水位预测

均方误差: 0.001545

均方根误差: 0.039309

平均绝对误差: 0.032946

图 4-40 LSTM 参数[4]48 小时水位预测误差

表 4-5 LSTM 神经网络连续预测误差对比

LSTM 参数/水位 误差	LSTM 参数[1] 学习率 0.05 步长 50	LSTM 参数[2] 学习率 0.01 步长 150	LSTM 参数[3] 学习率 0.005 步长 250	LSTM 参数[4] 学习率 0.001 步长 400
均方误差	0.059312	0.010267	0.003761	0.001545
均方根误差	0.243540	0.101328	0.061327	0.039309
平均绝对误差	0.209090	0.089104	0.050487	0.032946

从图 4-33~图 4-40 以及表 4-4 可知,在当前数据的情况下,学习率为 0.001,训练次数 400 次的情况下,模型对于短期水位预测效果较为优秀,但与参数[2]和参数[3]相差不大,均取得较好效果。

4.6 结论与应用

4.6.1 模型对比及结论

在 RNN 神经网络和 LSTM 神经网络中,在学习率为 0.001,训练次数为 400 次的情况下均取得较好的成果,为了更加直观的了解两种算法在最优情况下的模型,将图像进行整合,绘制图 4-41 和制作表 4-6,可以明显的看出 LSTM 神经网络对于真实的水位更加的接近,曲线更加拟合。并且各个误差参数都是 LSTM 神经网络更加的小,说明 LSTM 神经网络在苕溪流域的桥东村水位站水位预测更加准确。

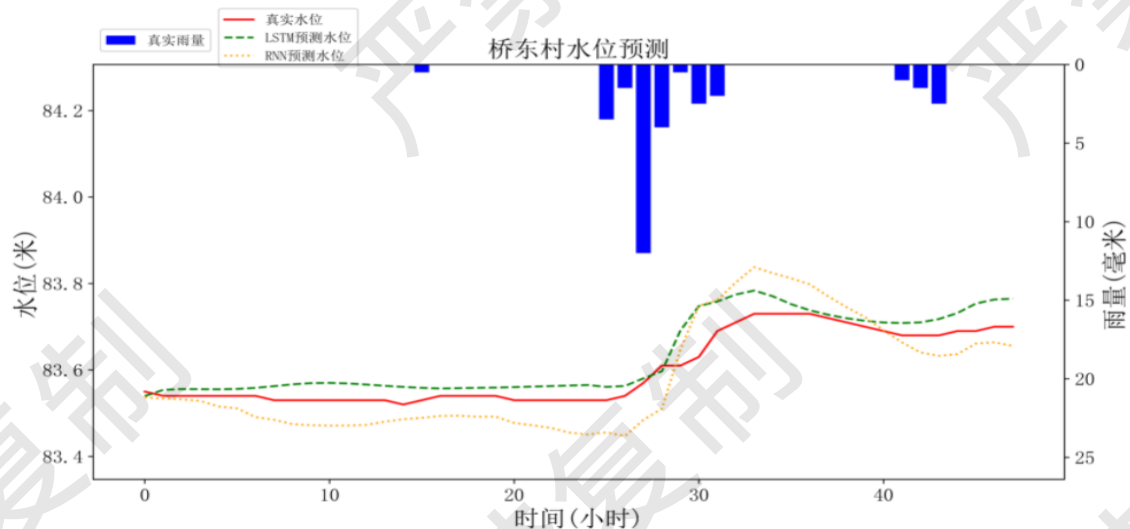


图 4-41 水位预测对比

表 4-6 水位误差对比

模型/水位误差	RNN 神经网络	LSTM 神经网络
均方误差	0.003473	0.001545
均方根误差	0.058928	0.039309
平均绝对误差	0.052403	0.032946

4.6.2 模型预测实战

根据上文的实现结论可以得出当学习率为 0.001, 训练次数为 400 次时的 LSTM 神经网络在茗溪小流域桥东村水位站的预测效果最佳。本小节浙江雨水情实时网站获取 2021 年 4 月 1 日至 12 日的历史水位信息, 从中央气象台获取 4 月 13 日的预测降水信息以及预测的空气湿度和温度信息。使用模型预测 4 月 13 日的水位, 预测结果如图 4-42 所示。

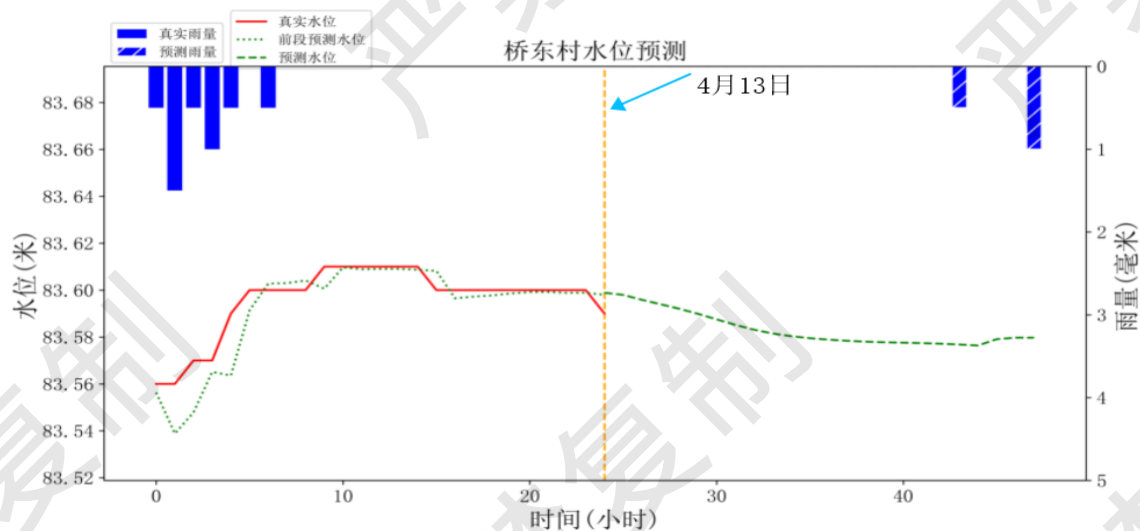


图 4-42 水位预测

根据 LSTM 模型预测的水位, 当预测过高时有关部门可以向群众发布洪水预警信息, 从而及时的疏散群众, 保障生命财产安全。

4.7 本章小结

本章阐述了实验数据的来源以及获取方法, 模型代码编写的流程, 以及编写代码过程中的细节, 并且分析了 RNN 神经网络和 LSTM 神经网络在不同的参数下模型的优劣。对未来 48 小时的水位进行了预测, 根据未来 48 小时水位的高低情况来判断是否发出预警信息, 应该发出何种等级的预警信息, 为应对山洪暴发提供了有理的保障。(1) RNN 神经网络和 LSTM 神经网络在最优参数的情况下, 模型训练均方误差分别为 0.000108 和 0.000064, 对近期水位进行预测均方误差分别为 0.003473 和 0.001545。(2) 在足够的训练次数和合适的参数情况下 LSTM 与 RNN 相近, 综合不同的模型来看, LSTM 神经网络的预测效果好于 RNN 神经网络。

结论

本文基于 RNN 神经网络算法和 LSTM 神经网络算法将浙江省临安市桥东村作为研究对象，并且构建山洪预警模型。以水位，雨量，空气湿度，温度作为输入指标，经过 RNN 神经网络算法和 LSTM 神经网络算法，输出水位信息，根据水位与警戒线以及保证线的距离来确定是否发布预警信息。

本文水位和雨量数据从浙江实时雨水情山洪预警平台获取，温度和空气湿度从 tutiempo 网站获取，在使用 python3.7 版本 tensorflow2.0 框架在 jupyter notebook 上进行编写算法，并且分析对比各个参数时算法的优劣性。实验表明：（1）RNN 神经网络和 LSTM 神经网络在最优参数的情况下，模型训练均方误差分别为 0.000108 和 0.000064，对近期水位进行预测均方误差分别为 0.003473 和 0.001545。（2）在足够的训练次数和合适的参数情况下 LSTM 与 RNN 相近，综合不同的模型来看，LSTM 神经网络的预测效果好于 RNN 神经网络。当模型预警的水位线超过桥东村预警水位线时可以发布预警信息，从而降低山洪灾害的损失。

限于本人时间和渠道有限，论文中还存在一些今后可以进一步解决的问题，包括如下方面：

（1）限于渠道有限，本文只能获取到近一年的数据，未来可以通过有关部门获取多年的数据，增加数据量，提高模型的精度。

（2）本文的数据爬取功能无法实时进行，还可以改进算法，增加实时爬取并自动保存清洗的功能，提高数据的实时性。

（3）可以增加小流域的站点，并且都训练出各自的模型，部署在网络上，方便山洪预测数据查看。

参考文献

- [1]叶勇. 浙江省小流域山洪灾害与预警技术研究[D]. 浙江大学, 2008.
- [2]王丹. 河南省典型小流域山洪灾害动态预警模型及应用[D]. 郑州大学, 2019.
- [3]. 水利部防御司召开 2020 年度全国山洪灾害防治项目建设管理工作会议[J]. 中国防汛抗旱, 2020, 30(11):4.
- [4]罗清元, 刘丽娜, 赵彦增, 赵延伟, 郭恒亮. 山洪灾害分析评价预警雨量应用机制与实践[J]. 水电能源科学, 2020, 38(10):52-55.
- [5]张帆. 基于人工智能的山洪灾害风险早期识别研究[D]. 华北水利水电大学, 2020.
- [6]史春朝. BP 神经网络算法的改进及其在 PID 控制中的应用研究[D]. 天津大学, 2006.
- [7]何秉顺, 郭良. 再谈山洪预警[J]. 中国防汛抗旱, 2018, 28(12):78-79+89.
- [8]陈培佳. 山洪灾害预警模型研究[D]. 浙江农林大学, 2017.
- [9]Moyang Liu, Yingchun Huang, Zhijia Li, Bingxing Tong, Zhentao Liu, Mingkun Sun, Feiqing Jiang, Hanchen Zhang. The Applicability of LSTM-KNN Model for Real-Time Flood Forecasting in Different Climate Zones in China[J]. Water, 2020, 12(2).
- [10]Boudani Mustapha Rabie, Mazour Mohamed, Mazighi Hichem, Djoukbala Omar. Development of a minimalist conceptual numerical model for flood forecasting and management under GIS environment[J]. Journal of Water and Climate Change, 2020, 11(S1).
- [11]M. Esfandiari, S. Jabari, H. McGrath, D. Coleman. FLOOD MAPPING USING RANDOM FOREST AND IDENTIFYING THE ESSENTIAL CONDITIONING FACTORS; A CASE STUDY IN FREDERICTON, NEW BRUNSWICK, CANADA[J]. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2020, V-3-2020.
- [12]Shusen Wang, Hazen A. J. Russell. Forecasting Snowmelt-Induced Flooding Using GRACE Satellite Data: A Case Study for the Red River Watershed[J]. Canadian Journal of Remote Sensing, 2016, 42(3).
- [13]Xuan-Hien Le, Hung Viet Ho, Giha Lee, Sungho Jung. Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting[J]. Water, 2019, 11(7).
- [14]吴昌友. 神经网络的研究及应用[D]. 东北农业大学, 2007.

致谢

本论文在汪松松导师的指导下完成。当我还在准备考研时，还在为毕业论文题目迷茫时，从对于神经网络，深度学习等概念极为模糊时，从汪松松老师那里了解到了人工智能领域的大致框架，有了一个可靠的路线。并且根据老师的指点选择了这个适合自己未来的题目。感谢老师我在不懂如何设计这个模型时提供明确的思路，在我撰写论文的路上给我精心的指点，论文中的许多要点都离不开您细心的指导。

时光荏苒，大学四年就要结束了。大二时当我羡慕同学可以靠自己的兼职等工作来实现自我独立时，我也希望能够不过分依赖父母的资助，于是通过靠游戏搬砖来赚钱，但是成绩却每况愈下。正当我迷茫时有一个人却通过自己的努力获得了一等奖学金和国家励志奖学金，这时我发现我错了，走上了一条歧路，但是为时不晚。当你问我要不要考研，我只是以试一试的心态参加考研，但因为相互激励，并且在考研的路上放弃了许多机会，使我们拥有了背水一战的气势，再苦再累也都挺过来，最终成功上岸。阮敬焱，感谢你。

感谢我的父母对我从小的养育之恩，感谢你们在我考研期间给了我最大的支持，做我最坚强的后盾，给我了一个安静的，可以全身心投入学习的环境。你们那忙忙碌碌的身影，那长满皱纹的双手，我知道，只有我不断向上，努力取得好成绩就是对您最好的礼物。在你们听到我的考研分数以及拟录取时，那时你们的开心是无法用言语来形容的。感谢你们对我无微不至的关怀。