

LSTM 在低能见度预报中的应用

高 鹏, 江 柯, 余涛涛

(中国民用航空飞行学院 空中交通管理学院, 四川 广汉 618307)

摘要:低能见度天气是制约交通出行的主要因素之一,提前做好低能见度的预报工作能利于提前决策、规避风险、降低损失。由于低能见度天气是时间序列问题,针对这种问题,近年来深度学习的长短时记忆网络模型能够显著提高预报精度,因此利用多变量数据对茅台机场低能见度进行预测。首先进行气象因素分析,利用皮尔逊相关系数法来挑选相关性高的气象因子,从而减小数据量得到更利于训练的数据集。然后利用 LSTM 网络对多元时间序列进行建模,实现机场低能见度的预测。经过试验,模型对茅台机场能见度的准确预报率为 85.43%,为机场低能见度预报提供了新的方法。

关键词:多元时间序列;长短时记忆网络;低能见度;机场

中图分类号:TP183;TP391.4 **文献标志码:**A **文章编号:**1671-1807(2023)03-0273-06

近年来,随着航空运输的迅速发展,企业和用户对运输准点率的要求也越来越高,而天气因素是降低航班准点率的一大诱因,其中大雾天气在其中占比很大。设备落后的中小型机场一旦遭遇大雾造成的低能见度天气,面临的将是航班的返航与备降。因此提高大雾的预报准确率才能利于机场用户提前决策,降低低能见度天气带来的损失。

大雾的预报在过去常以人工主观预报为主,受人工误差影响较大,而近年来的数值模式预报、动力统计预报以及集合预报等手段都能显著提升预报准确率。伴随着计算机技术的不断进步,使得人工智能中的神经网络模型在利用海量历史数据对气象要素的预报中大展身手。通过分析能见度历史样本的统计规律,建立能见度样本之间的非线性关系^[1]。这类方法包括人工神经网络(artificial neural network, ANN)、支持向量机(support vector machine, SVM)等人工智能预测方法^[2-3]。但是这类方法外推预报需要消耗大量的计算机资源,且对能见度物理机制掌握不足,因此比较适用于短期和超短期预测任务^[4]。

卷积神经网络通过视频帧序列图像建立能见度与图像输入数据的非线性关系,通过卷积层提取图像特征,分析数据特征与能见度的关系,然后将训练好的模型用于能见度预测。卷积神经网络对

于视频、图像处理能力比较强,但对于序列数据前后关联性不强^[5]。因此提出了循环神经网络,能够考虑时间序列的时序相关性,理论上能够利用任意长度的历史信息来更加全面完整地对时间序列进行建模。作为一种特殊的循环神经网络(recurrent neural network, RNN)模型^[4],长短期记忆(long short-term memory, LSTM)网络^[6]通过自身特殊的结构设计,有效地规避了常规 RNN 训练过程中带来的梯度消失和梯度爆炸问题^[7-8],使得数据能够有效地被训练^[9],从而提升能见度预测效果。

由此本文基于长短时记忆网络(LSTM)模型,利用茅台机场海量高精度历史数据对茅台机场大雾能见度进行预报,实验表明更为复杂的长短时记忆网络能够有效提高茅台机场低能见度预测精度。

1 LSTM 的构建及预报应用

1.1 LSTM 模型简介

LSTM 是一种特殊的 RNN。一般的 RNN 展开方式是将几个功能相同的单元连续串联,即随着新输入信息的数量进一步扩大,通过上述自我循环的过程将上一个的状态信息传送给当前的,同时再对新的输入信息进行当前轮次的训练或者学习,一直到整个系统的训练全部完成,而最后获得的数据即是系统最后的预测结果。普通的 RNN 模型如图 1^[9]所示,图中 o_t 表示输出层在 t 时刻的输出; s_t

收稿日期:2022-09-04

作者简介:高鹏(1997—),男,四川成都人,中国民用航空飞行学院空中交通管理学院,硕士研究生,研究方向为空中运行环境。

表示隐藏层在 t 时刻隐藏状态值; x_t 表示 t 时刻的输入值; V 表示隐藏层到输出层的权值矩阵; W 表示隐藏层到隐藏层的权重矩阵; U 表示输入层到隐藏层的权重矩阵。

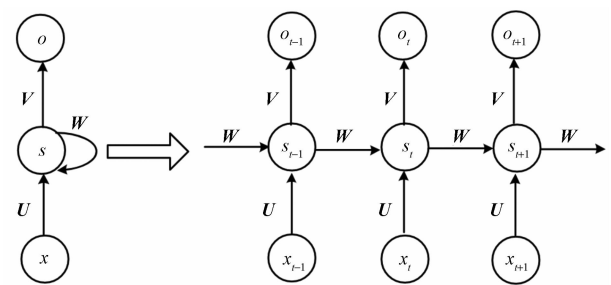


图 1 普通 RNN 模型

原始的 RNN 系统在训练时,由于训练长度的增加或者网络层数的增加,很容易产生梯度弥散爆炸和局部梯度消失的现象,使得系统无法处理更长的信息,也因此无法得到更长距离数据的信息。LSTM 的出现解决了在递归式神经网络中普遍存在的长期依赖现象,它通过内部控制信息传递的“门”一定程度上避免了梯度消失和梯度爆炸的出现^[7-8]。最近,LSTM 网络已经在视频分类^[11]、语音理解^[12]、文本翻译上得到应用。

LSTM 是 RNN 模型演化出来的变体,两者的区别在于普通 RNN 结构简单,隐藏信息只源于当前时刻输入和上一时刻的隐藏层信息,没有记忆功能。LSTM 网络结构复杂,不仅引入了细胞状态,而且使用输入门、遗忘门、输出门 3 种门来保持和控制信息。LSTM 网络结构如图 2^[10]所示,图中 i 、 f 、 o 、 c 、 h 分别表示输入门、忘记门、输出门、内部状态、单元输出; W 表示权重矩阵; b 表示偏置项; σ 表示 sigmoid 函数; $\tanh$ 表示双曲正切激活函数; $\odot$ 表示哈达马乘积。

LSTM 细胞由输入门、忘记门、输出门和内部状态构成。输入门是判断当前时刻系统的输入输出信息中有哪些需要存储的候选状态;忘记门是控制上一时刻的单位状态,有哪些必须需要被遗忘;而输出门则是控制当前单位状态有哪些必须输出到当前的输出值。

1.2 LSTM 的训练方法

目前,针对 LSTM 等递归神经网络模型,主流的训练方法有两种,按时间展开的反向误差传播算法(BPTT)^[8]和实时递归学习算法(RTRL)^[13]。由于 BPTT 算法概念清晰且计算高效,在计算时间上较 RTRL 更具优势,因此本文采用该算法来训练

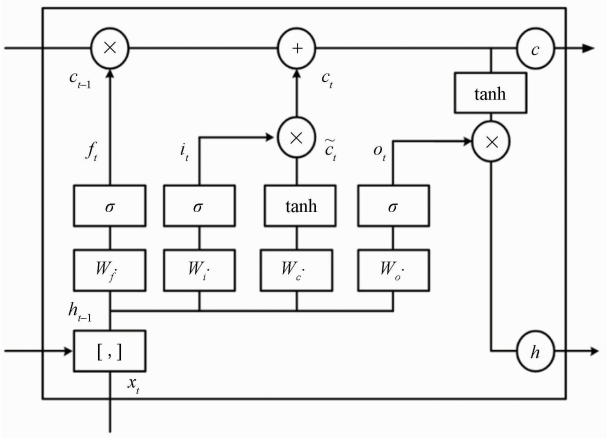


图 2 LSTM 网络结构

LSTM 网络。

对于 LSTM 来说有一个时序概念,就是需要把梯度沿时间通道传递的时间 BP 算法,所以称之为 Back Propagation Through Time-BPTT。BPTT 基本思路就是首先把整个 LSTM 网络按时间顺序展开,之后再通过时间误差反向传播计算对扩展后的网络进行时间训练。与标准的 BP 算法相似,BPTT 也是需要重复使用链式规则。与标准前馈神经网络不同的是,对于 LSTM 网络,损失函数不但与输入层次状态相关,也同时与上一隐藏层的状态相关,且 LSTM 每一级的权重都是共享的。

2 机场能见度预测模型设计

2.1 数据说明

本文使用的 3 年(2017—2019 年)茅台机场大雾数据来自机场 35 号跑道的自动气象观测系统(AWOS)1 min 4 次的能见度 VIS、风速数据 WIND SPEED,1 min 1 次的湿度 HUMIDITY、温度 TEMP、机场修正海平面气压数据 QNH,其时间分辨率较高。

2.2 数据预处理

由于茅台机场于 2017 年正式开航且位于山区,因此受气象条件以及当地环境等因素影响容易导致数据的缺测、异常。因此本文前期针对数据进行了大量处理工作。

本文利用前向填充和均值平滑法对数据进行缺省值处理。对于 1 min 4 次的时间分辨率较高的数据,采用前向填充法即用前一值代替后面的缺测值;而对于 1 min 1 次的的数据,采用均值平滑法,即用临近时间的均值来代替中间的缺省值。

在使用多变量时序序列进行大雾能见度预测时,由于量纲的不同,数值差异较大,例如 24 h 降水

和气压值量纲相差较大。为避免数值过小的数据被忽略,数值过大的数据对权重占比大,同时为加快模型收敛,需要对变量和能见度值进行归一化处理^[14]。其归一化公式为

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

式中: x' 是数据进行归一化后的结果; x 为气象初始数据; $x_{\max}$ 为每种气象要素对应的最大值; $x_{\min}$ 为每种气象要素对应的最小值。

特别地,针对风向,考虑其物理意义,使用正弦函数对其进行归一化,将其数值归算到区间 $[-1, 1]$ 内^[1]。

2.3 机场能见度与气象因素相关性分析

机场大气能见度数值是局地变化的,其时间精度和空间精度较小,周围小的扰动便会引起能见度数值发生波动,因此只基于历史能见度数据对当前能见度的短时预测方法存在着明显的不足,而研究各个气象要素对能见度的影响,并基于皮尔逊相关系数法筛选预报要素因子,既能提高外推潜力,又能够有效降低预测模型的噪声。

将经过数据处理后的 AWOS 数据利用 Pearson 相关系数法公式进行计算,其中 Cov 表示协方差, σ 表示标准差, μ_X 表示 X 的期望, μ_Y 表示 Y 的期望。其 Pearson 相关系数的绝对值越接近于 1,则代表数据之间的相关性越好,而越接近于 0,则代表相关性越差。负值代表数据间呈现出负相关关系,正值代表数据间呈现出正相关关系。

$$\rho_{X,Y} = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \tag{2}$$

代入经人工经验判别的影响机场能见度的 19 个气象要素数据,计算可得温度和湿度、10 min 平均风速与能见度的相关性最大,如图 3 所示,3 年平均值分别为 0.621、-0.574 2、0.587 6。气压、24 h 降水与能见度的相关性次之,3 年平均值分别为 -0.373 8、-0.037。由此可见,在预报能见度值时的温压湿降水是其主要影响因素,因此采用将影响能见度最显著的 6 组气象数据序列与能见度数值序列构成多变量的序列来使用多个气象要素序列预测能见度值。

2.4 LSTM 预测模型参数

构建机场低能见度短期预测的 LSTM 模型,主要需要确定模型的 5 个超参数,即输入层时间步长、输入层维数、隐藏层层数、隐藏层节点个数以及输

出层维数。

输入层时间步长是表示用于进行机场低能见度预测的变量时间序列的长度。该参数的设定既要考虑到预报信息的完整性,也要兼顾到模型训练的实效性。过长的历史序列长度会导致模型复杂度较高,可能导致模型的过拟合,从而得到恶化模型;过短会缺失预测信息导致模型预测精度难以提高^[15-16]。本文通过模拟网络优化算法,将该参数设置为 18,即输入前 18 个时刻(前 3 h)的历史数据用于预测;输入层维数即变量数,在单变量预测时,输入层维数为 1;隐藏层层数就是 LSTM 层的个数,隐藏层数越多学习到的经验越多,即模型的非线性拟合能力提高,但同时模型的复杂度及计算成本也将增加。针对大雾能见度预测,本文通过模型优化,将该参数设为 1;同样的方法将隐藏层节点个数设为 128;由于本次预测任务是预测大雾能见度的具体数值输入回归任务,因此将输出层维数设为 1。

3 模型预测效果分析

3.1 数据说明和实验设计

本文采用的样本数据为 AWOS 系统提供的茅台机场 2017—2019 年连续 3 年的实测数据,获取的数据包括风速、能见度、机场修正海平面气压数据、温度、湿度以及 24 h 降水量等 6 种不同的时间序列,采样的间隔为 10 min,总计 154 961 条样本数据。

3.2 预测性能分析

利用输入的前 18 个时刻的 LSTM 模型预测后 10 min、1 h、2 h 的能见度。利用平均绝对误差(mean absolute error, MAE)来评价此次模型的误差和准确度。MAE 的公式为

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - y| \tag{3}$$

式中: y_i 表示预测值; y 表示实测值。MAE 越大,则代表误差越大,模型的预测效果越差。

输入的前 18 个时刻的 LSTM 模型预测后 10 min 的损失函数曲线、能见度实测值和预测值拟合曲线如图 4 所示。

由图 4 可见该模型此次预测的效果较好,预报准确率为 85.43%,预测曲线和真实曲线的拟合程度高,MAE 为 2 146.78 m。

输入的前 18 个时刻的 LSTM 模型预测后 1 h 的损失函数曲线、能见度实测值和预测值曲线如图 5 所示。

由图 5 可见该模型此次预测的效果较好,但没

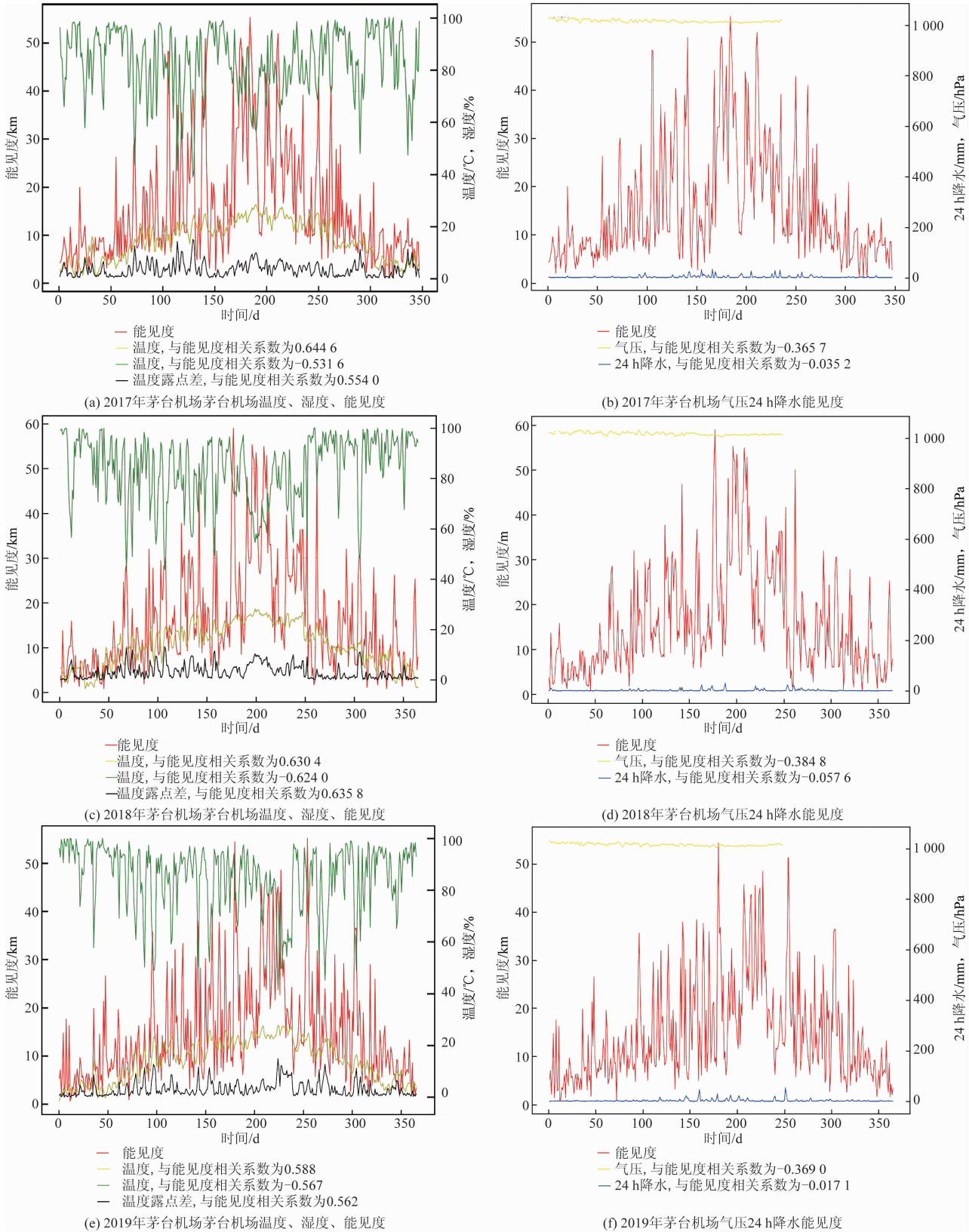


图 3 2017—2019 年茅台机场温度、湿度、能见度折线图和气压 24 h 降水能见度折线图

有利用 3 h 预测后 10 min 的高,预报准确率为 83.97%,预测曲线和真实曲线的拟合程度仍然较

高,MAE 为 2 367.55 m。
输入的前 18 个时刻的 LSTM 模型预测后 2 h

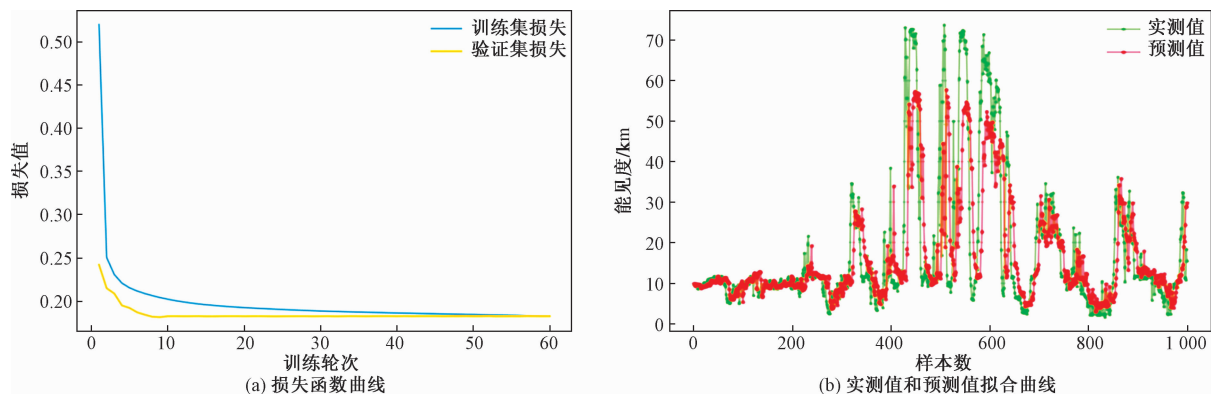


图 4 预测 10 min 的损失函数曲线、能见度实测值和预测值拟合曲线

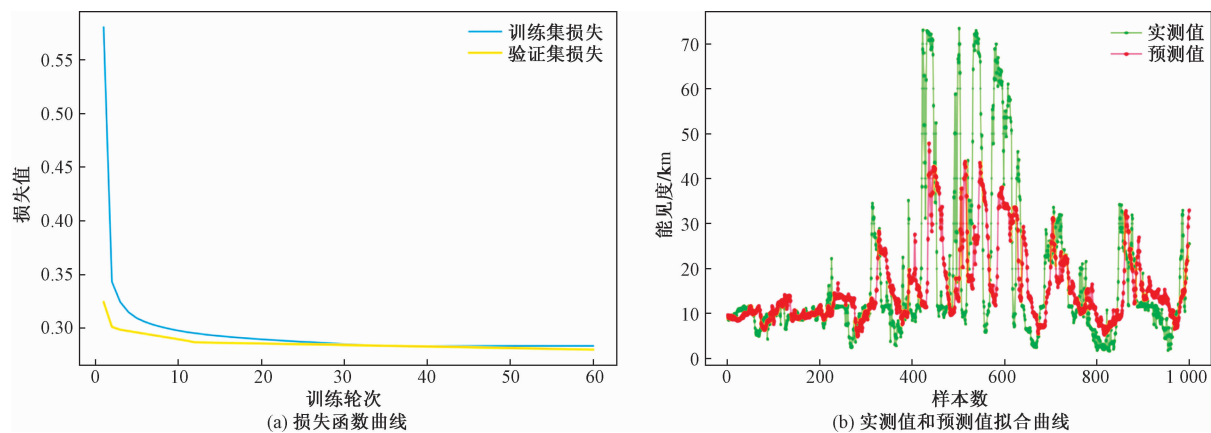


图 5 预测后 1 h 的损失函数曲线、能见度实测值和预测值拟合曲线

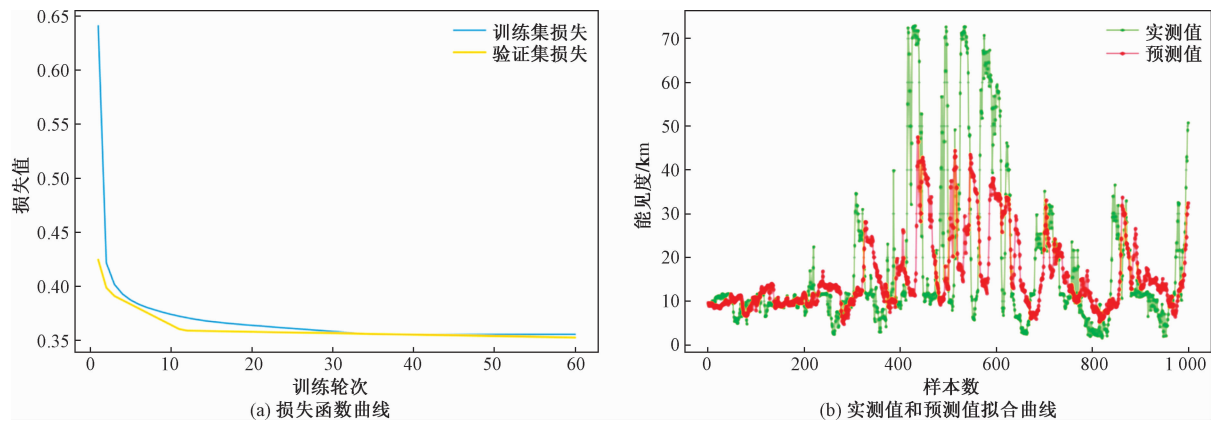


图 6 预测 2 h 后的损失函数曲线、能见度实测值和预测值拟合曲线

的损失函数曲线、能见度实测值和预测值拟合曲线如图 6 所示。

由图 6 可见该模型此次预测的效果是三者中预测效果最差的,预报准确率为 80.22%,但仍然较为精准,预测曲线和真实曲线的拟合程度良好,MAE 为 2 421.92 m。可见随着预测时效的提高,预报效果逐渐变差,且其滞后性水平下降。

4 结论

本文提出了基于 LSTM 的深度学习网络的低能见度客观预报方法,根据实验结果表明预报准确率为 85.43%,因此可以很好地提升低能见度准确预报率,更好地保障茅台机场航班的正常运行。精细化的低能见度人工智能预报系统对推动主观预报向客观预报发展,从定性走向定量,提高预报精

准性,具有重要意义。通过此次预报分析发现温度、湿度对于机场能见度的预报更重要。并且通过单步滚动预测的方案做对比实验发现,时间步长越长效果越差。下一步可以考虑利用 transformer 做长序列预测。

参考文献

[1] 李智,韩学山,韩力,等. 地区电网风电场功率超短期预测方法[J]. 电力系统自动化,2010,34(7):90-94.

[2] 许可. 卷积神经网络在图像识别上的应用的研究[D]. 杭州:浙江大学,2012.

[3] 茆美琴,周松林,苏建徽. 基于脊波神经网络的短期风电功率预测[J]. 电力系统自动化,2011,35(7):70-74.

[4] 朱乔木,李弘毅,王子琪,等. 基于长短期记忆网络的风电场发电功率超短期预测[J]. 电网技术,2017,41(12):3797-3802.

[5] 黄晋,李保强,吕明燕,等. 基于 VGG 卷积神经网络的机场能见度预测[J]. 集成电路应用,2022,39(3):58-59.

[6] HOCHREITER S,SCHMUDHUBER J. Long short-term memory[J]. Neural Computation,1997,9(8):1735-1780.

[7] HOCHREITER S,BENGIO Y,FRASCONI P,et al. Gradient flow in recurrent nets; the difficulty of learning long-term dependencies[C]//A Field Guide to Dynamical Recurrent Networks, Wiley-IEEE Press,2001:237-243.

[8] BENGIO Y,SIMARD P,FRASCONI P. Learning long-term dependencies with gradient descent is difficult[J].

IEEE Transactions on Neural Networks,1994,5(2):157-166.

[9] TIAN Y,LI P. Predicting short-term traffic flow by long short-term memory recurrent neural network[C]//IEEE International Conference on Smart City, Chengdu: IEEE, 2015.

[10] 徐高扬. 基于循环神经网络的台风路径预测模型研究[D]. 成都:西南交通大学,2019.

[11] GAO L,GUO Z,ZHANG H,et al. Video captioning with attention-based LSTM and semantic consistency[J]. IEEE Transactions on Multimedia,2017,19(9):2045-2055.

[12] SONG E,SOONG F K,KANG H G. Effective spectral and excitation modeling techniques for LSTM-RNN-Based speech synthesis systems[J]. IEEE/ACM Transactions on Audio, Speech and Language Processing, 2017,25(11):2152-2161.

[13] CHANG F,CHANG C,HUANG H,et al. Real-time recurrent learning network for stream-flow forecasting[J]. Hydrological Processes,2002,16(13):2577-2588.

[14] 李素萍,姚洪兴. 多变量时间序列分析及其应用[J]. 科学技术与工程,2008(11):2764-2769.

[15] 王愈轩,梁沁雯,章思远,等. 基于 LSTM-XGboost 组合的超短期风电功率预测方法[J]. 科学技术与工程,2022,22(14):5629-5635.

[16] 韩朋,张晓琳,张飞,等. 基于 AM-LSTM 模型的超短期风电功率预测[J]. 科学技术与工程,2020,20(21):8594-8600.

Application of LSTM in Low Visibility Forecasting

GAO Peng, JIANG Ke, YU Taotao

(School of Air Traffic Management,Civil Aviation Flight University of China,Guanghan Sichuan 618307,China)

Abstract: Low visibility weather is one of the main factors restricting traffic travel,so forecasting low visibility in advance can help make decisions in advance,avoid risks and reduce losses. Since low visibility weather is a time series issues,the long and short-term memory network of deep learning can significantly improve the forecast accuracy in recent years, multivariate data is used to predict the low visibility of Maotai Airport. Correlation analysis, using Pearson similarity to screen for factors with high correlation,thereby reducing the complexity of the data. The LSTM network is then used to model multivariate time series to predict airport low visibility. After experiments,the model’s accuate forecast rate of low visibility at Maotai Airport was 85.43%,which provided a new forecasting method for airport for low visibility.

Keywords: multivariate time series;long short-term memory;low visibility;aerodrome