

# 基于 LSTM 的机场跑道视程预测<sup>①</sup>



彭路, 柳俊凯, 盛爱晶, 张兴海, 孙文正

(南京恩瑞特实业有限公司, 南京 211106)

通信作者: 彭路, E-mail: 812234987@qq.com; 柳俊凯, E-mail: liujunkai2013@outlook.com

**摘要:** 跑道视程反映了飞行员的视程大小, 是保障飞机起飞和着陆飞行安全的重要气象要素之一, 提升跑道视程的预测准确率将有效提升飞机在低能见度和复杂天气条件下的起降能力和航空管制能力. 跑道视程除了受自身观测仪器不足限制, 雾、烟、沙尘、强降水等其他天气对其影响也十分显著. 本文利用 2012–2018 年咸阳机场民航自动气象观测系统观测的风速、湿度、温度和跑道能见度等气象要素时间序列数据, 首先分析跑道视程与其他气象要素之间的长期相关关系, 并基于相关分析结果采用人工智能领域最常用的长短时记忆网络, 构建了一种机场跑道视程预测模型, 模型实验结果表明该模型在 0–2 小时内跑道视程预测平均拟合度能够达到 72%.

**关键词:** 长短时记忆网络; 深度学习; 跑道视程预测; 时序预测; 神经网络; 预测模型; 人工智能

引用格式: 彭路, 柳俊凯, 盛爱晶, 张兴海, 孙文正. 基于 LSTM 的机场跑道视程预测. 计算机系统应用, 2022, 31(5): 203–212. <http://www.c-s-a.org.cn/1003-3254/8492.html>

## Prediction of Runway Visual Range Based on LSTM

PENG Lu, LIU Jun-Kai, SHENG Ai-Jing, ZHANG Xing-Hai, SUN Wen-Zheng

(Nanjing NRIET Industrial Co. Ltd., Nanjing 211106, China)

**Abstract:** Runway visual range (RVR) reflects the pilot's visual range, which is one of the important meteorological elements to ensure aircraft flight safety when the aircrafts take off and land. Improving the prediction accuracy of RVR will effectively improve the aircraft's take-off and landing ability and aviation control ability under low visibility and complex weather conditions. The RVR is mainly affected by fog, smoke, dust, heavy precipitation and other weather, as well as the lack of instruments. According to the time series data of meteorological elements such as wind speed, humidity, temperature and runway visibility observed by the civil aviation automated weather observing system of Xianyang Airport from 2012 to 2018, this study firstly analyzes the long-term correlation relationship between the RVR and other meteorological observation data. On the basis of the correlation analysis, this study also uses the long short term memory network (LSTM), which is the most commonly used in artificial intelligence field, to construct an airport RVR prediction model. The experiment results show that the average fitting degree of the model can reach 72% within 0–2 h.

**Key words:** long short term memory network (LSTM); deep learning; prediction of runway visual range; time series prediction; neural network; prediction model; artificial intelligence

跑道视程 (runway visual range, RVR) 是指起飞/着陆时飞机与跑道内地面物体之间的视程, 表示在大雾等恶劣天气条件下跑道和跑道中心线的识别程度,

是影响飞机起飞和降落的重要因素之一. 在能见度小于 800 m 的情况下, 航空管制员需要根据跑道视程判断飞机是否能够起降<sup>[1]</sup>. 随着民航航班总量的快速增

① 收稿时间: 2021-07-15; 修改时间: 2021-08-18, 2021-09-18; 采用时间: 2021-09-29; csa 在线出版时间: 2022-04-11

长,跑道视程对航班运行的影响愈加显著,已经成为影响航班安全、正常、效率的主要因素<sup>[2]</sup>。实际机场运行管制过程中,为了保障起降安全,一般较低的跑道视程发生时不允许飞机起降。相应地,跑道视程降低将导致机场运行能力减弱,机场航班容量降低,进而引起航班堵塞延误,旅客滞留,增加机场运行成本。因此,机场气象或者管制人员就会特别关心未来一段时间内跑道视程的变化趋势,跑道视程预测将能够为机场业务人员提前决策提供依据,特别是对低跑道视程条件提前作出预警,制定相关起降、管制预警应急方案。

跑道视程不是一个直接观测量,目前许多机场已经布设了民航自动气象观测系统(automated weather observing system, AWOS)通过参数计算实时的跑道视程值,但观测仅仅能够反映当前状态,无法满足机场业务部门对未来跑道视程变化趋势的产品需求,机场业务上至少需要提前预报未来近6个小时的跑道视程变化情况辅助业务决策。

机场跑道视程预测常用的传统方法多是基于不同天气形势的各种预报模型的知识积累和近似记忆,但因知识更新手段的繁琐,近年来该方法逐渐被预报员抛弃<sup>[3]</sup>。随着计算机计算能力以及人工智能技术的进步,深度学习预测技术在跑道视程预测领域日益受到关注。朱蕾等人利用3年的乌鲁木齐机场RVR观测资料,使用SVM方法分别构建了大于800m和小于800m的RVR推理模型<sup>[3]</sup>。Ku等人也采用了卷积神经网络对韩国泰安机场的跑道视程预测进行了初步尝试<sup>[4]</sup>。周志洋利用数据挖掘技术,通过多个维度的气象特征首次对高原机场的跑道视程进行了短时预测,准确率最高达到81.14%<sup>[5]</sup>。

长短时记忆网络(long short term memory network, LSTM)是一种改进的循环神经网络,它有效解决了循环神经网络在梯度消失和缺乏长期记忆能力等方面存在的问题,能够从海量的长时间序列信息中自动学习其特征并保存,适合处理时序预测问题,近年来在气象领域已得到应用。袁芳介绍了LSTM技术的基本原理,并提出该方法在处理气象领域非线性分类、回归等问题应用建议<sup>[6]</sup>。李芬等人将LSTM方法用于细颗粒物(particle matter $\leq 2.5\ \mu\text{m}$ , PM<sub>2.5</sub>)的短期预测<sup>[7]</sup>。魏昱洲等人使用LSTM方法进行超短期风速的预测试验,均取得了一定的预报效果<sup>[8]</sup>。此外,罗芳琼等人基于LSTM方法提取的特征信息,结合随机森林和支持向量机方

法,建立了低温气象灾害预报混合模型,取得了较高的预测精度<sup>[9]</sup>。

低跑道视程现象受多重气象和非气象因子影响,不同因子之间也同样存在相互作用关系,具有明显的非线性特征,预报难度较大。因而,如何尽可能地提高对低跑道视程天气的预报能力,从而实现提前预警是航空气象保障人员颇为关注的问题。为确保飞机能够安全起降,本文采用咸阳机场2012–2018年共7年的民航自动气象观测系统观测资料对跑道视程进行研究,采用新型的基于长短时记忆网络方法,建立机场跑道视程预测模型,旨在为气象保障人员提供一定的可供参考的预测产品,辅助决策保障飞机的安全起降。

## 1 跑道视程、长短时记忆网络概念解析

### 1.1 跑道视程

跑道视程是指在机场跑道中线上,飞机上的飞行员能够看到跑道面上的标志或者跑道边界灯或中线灯的距离<sup>[10]</sup>。跑道视程不是直接测量的气象要素,通过跑道灯光强度( $I$ )、照度阈值( $E$ )和大气消光系数( $\sigma$ ),根据柯西米德定律(Koschmieder's law)和阿拉德定律(Allard's law)计算得出的数值<sup>[2]</sup>,将两个定律计算得出的较大值作为要报告的RVR值。

在白天无灯光的情况下,根据柯西米德定律,RVR的计算公式为<sup>[2]</sup>:

$$RVR = -\ln 0.05/\sigma \approx 3/\sigma \quad (1)$$

在白天能见度受限或夜晚的情况下,此时需要开启跑道灯,依据阿拉德定律,RVR的计算公式为<sup>[2]</sup>:

$$\sigma RVR + \ln(E) - \ln(I) = -2 \ln(RVR) \quad (2)$$

其中,跑道灯光强度可通过背景亮度仪测量的背景亮度估算得出,照度阈值可按照设定的跑道灯光级数估算得到,大气消光系数可由大气透射仪或前向散射仪测量的气象光学视程得出<sup>[2]</sup>。

从公式出发,跑道视程似乎与气象无关,但实际业务中,照度阈值及大气消光系数的变化受到多种气象要素综合影响,例如强降水导致能见度降低,直观上导致大气消光系数增大,跑道视程降低,大雾也能导致能见度降低,直观上导致大气消光系数增大,跑道视程降低<sup>[2]</sup>。跑道灯光级数调整导致照度阈值变化也受降水、大雾等低能见度天气的影响。气象要素对跑道视程的影响作用是极其复杂的、非线性的,还没有任何一种数学物理公式或者模型能够完全刻画这种影响关系。

咸阳机场 I 类精密进近着陆标准为主导能见度  $\geq 800$  m 或  $RVR \geq 550$  m, II 类精密进近着陆标准为  $300 \text{ m} \leq RVR < 550 \text{ m}$ <sup>[11]</sup>. 对于 II 类精密进近, 当跑道视程观测设备故障或无法获得跑道视程报告时不得实施 II 类运行<sup>[11]</sup>. 通常, 将  $RVR < 550$  m 定义为低跑道视程, 低跑道视程下常引发大面积航班延误, 严重影响航班正常运行, 易引发不安全事件, 飞行员及各安全保障部门均需要非常谨慎地对待这种运行情况<sup>[1,11]</sup>.

## 1.2 长短时记忆网络

LSTM 是 Hochreiter 等人<sup>[12]</sup> 1997 年提出的基于循环神经网络理论的具有记忆长短期信息能力的神经网络, 是一种特殊的循环神经网络 (recurrent neural network, RNN), 主要是为了解决长序列训练过程中的梯度消失和梯度爆炸问题, 相比普通的 RNN, LSTM 在更长的序列中能够有更好的表现<sup>[7,13]</sup>. 经过若干代的发展, 已经形成了比较系统且完整的 LSTM 框架, 并且在语音识别、图像描述、自然语言处理、市场调研和交通预测等领域得到了广泛的应用<sup>[14-16]</sup>. LSTM 的基本节点结构如图 1 所示.

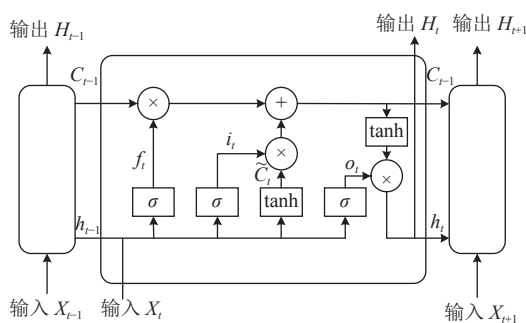


图 1 LSTM 基本节点结构

图 1 中,  $\times$  表示缩放信息,  $+$  表示添加信息,  $\sigma$  表示 Sigmoid 层,  $\tanh$  表示 tanh 层 (双曲正切),  $H_{t-1}$  表示上一个节点的输出,  $C_{t-1}$  表示上一个节点的状态,  $X_t$  表示输入,  $C_t$  表示更新后的状态,  $H_t$  表示输出. 贯穿图 1 的两条水平线代表了传输单元的状态, 通过线性变换保证了单元内部信息传输的不变性, 它是长短时记忆网络的核心<sup>[16]</sup>. 长短时记忆网络还能通过几个结构管理信息的传输, 从而添加或删除向单元状态中传输的信息, 选择性的让信息通过, 一般将这些结构称为门限<sup>[17]</sup>. LSTM 内部主要有 3 个阶段 (遗忘、输入、输出), 这 3 个阶段通过 3 个门限逻辑实现, “遗忘门”忘记部分过去的信息, “输入门”记住部分现在的信息, 然后将过去

的记忆与现在的记忆合并后通过“输出门”决定最终输出的部分.

(1) 遗忘阶段: 这个阶段主要是对上一个节点传进来的输入进行选择性的忘记, 遗忘门计算逻辑:

$$f_t = \sigma(W_f \cdot [H_{t-1}, X_t] + b_f) \quad (3)$$

其中,  $f_t$  表示遗忘门输出,  $H_{t-1}$  表示上一个节点的输出,  $X_t$  表示输入,  $W_f$  表示遗忘门权重,  $b_f$  表示遗忘门偏置,  $\sigma$  表示 Sigmoid 层.

(2) 输入阶段: 这个阶段将这个节点的输入有选择性的进行“记忆”, 输入门计算逻辑:

$$i_t = \sigma(W_i \cdot [H_{t-1}, X_t] + b_i) \quad (4)$$

其中,  $i_t$  表示输入门输出,  $H_{t-1}$  表示上一个节点的输出,  $X_t$  表示输入,  $W_i$  表示输入门权重,  $b_i$  表示输入门偏置,  $\sigma$  表示 Sigmoid 层.

当前候选节点状态计算逻辑:

$$\tilde{C}_t = \tanh(W_C \cdot [H_{t-1}, X_t] + b_C) \quad (5)$$

其中,  $\tilde{C}_t$  表示当前候选节点状态,  $H_{t-1}$  表示上一个节点的输出,  $X_t$  表示输入,  $W_C$  表示权重,  $b_C$  表示偏置,  $\tanh$  表示 tanh 层.

将前两个阶段得到的结果通过计算逻辑:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (6)$$

其中,  $C_t$  表示传输给下一个节点的状态,  $C_{t-1}$  表示上一个节点的状态.

(3) 输出阶段: 这个阶段将决定哪些信息将会被当成当前状态的输出, 输出门计算逻辑:

$$o_t = \sigma(W_o \cdot [H_{t-1}, X_t] + b_o) \quad (7)$$

其中,  $o_t$  表示输出门输出,  $H_{t-1}$  表示上一个节点的输出,  $X_t$  表示输入,  $W_o$  表示输出门权重,  $b_o$  表示输出门偏置,  $\sigma$  表示 Sigmoid 层.

当前节点最终输出还对上一阶段得到的  $C_t$ , 通过 tanh 函数进行了放缩, 计算逻辑:

$$H_t = o_t \cdot \tanh(C_t) \quad (8)$$

其中,  $H_t$  表示当前节点的最终输出.

## 2 跑道视程预测数据集构建

### 2.1 AWOS 观测资料

本文采用了 2012–2018 年咸阳机场共 7 年的民航气象自动观测系统 AWOS 观测资料, AWOS 每分钟记录一次各种气象要素数据, 包括风向、风速、温度、

露点温度、气压、相对湿度、降水量、背景亮度、气象光学视程、跑道视程等。

咸阳机场南北两条跑道沿东北-西南方向, AWOS 观测设备位于两条跑道的 4 个跑道端点及 2 个中点, 南侧跑道东北端点编号 RWY23L, 中点编号 SOUTH MID, 西南端点编号 RWY05R, 北侧跑道东北端点编号 RWY23R, 中点编号 NORTH MID, 西南端点编号 RWY05L。

## 2.2 预报因子选择

影响机场跑道视程的天气主要有雾、烟、沙尘、强降水等, 不同的天气现象具有不同的气象影响因子。因此, 本小节主要通过对比跑道视程的时序特征分析和相关分析, 获取影响咸阳机场的典型气象因子。

通过对咸阳机场 2012-2018 年共 7 年的 AWOS 观测资料统计, 咸阳机场各端点低跑道视程的月际分布如图 2 所示。

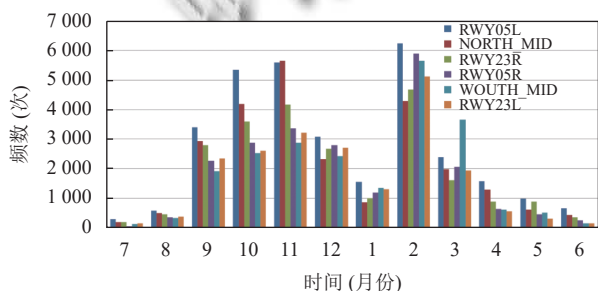


图2 咸阳机场各端点低跑道视程月际分布图

图2中低跑道视程天气的月际变化呈双峰型分布, 集中多发于秋冬季的9-12月和2-3月, 春季和夏季发生次数较少, 这种季节分布特征表明本场低跑道视程天气主要是由雾造成。而形成雾的相关气象因子条件有: 地面气温相对较低, 大气相对湿度较大, 接近准饱和状态, 近地层无风。

经过统计, 本场各端点低跑道视程的小时分布如图3所示。

图3中低跑道视程天气多发于凌晨至上午, 下午至傍晚较少发生, 并且在凌晨至清晨北侧跑道明显多于南侧跑道, 而上午发生的低跑道视程天气南侧跑道多于北侧跑道。表明咸阳机场的大雾天气主要以辐射雾为主, 机场西北部平原相对东南方向的渭河河谷辐射冷却效应更强, 夜间至凌晨更容易产生辐射雾, 而日出后北部升温快于南部, 辐射雾消散更快, 另一方面和渭河河谷冷气垫的温差造成的局地环流利于河谷区域

的雾自东南向西北方向输送。

此外, 针对 AWOS 自动观测资料中存在气象要素较多, 以及考虑到跑道视程也与仪器自身特性相关。为了尽可能多地获取与低跑道视程相关的各类因子, 利用本场 AWOS 自动观测每分钟的平均风、阵风、温度、露点温度、气压、相对湿度、背景光亮度等地面观测要素进行相关性分析, 从而选择预报因子。

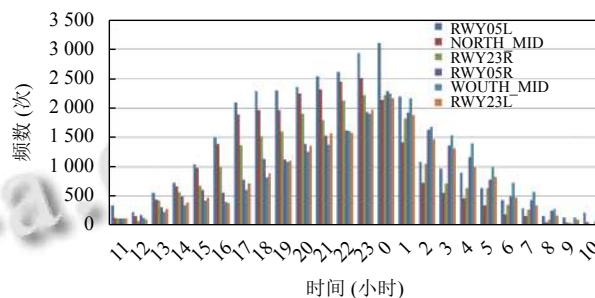


图3 咸阳机场各端点低跑道视程小时分布图

AWOS 各地面观测要素之间的相互相关性特征如图4所示。

图4可见, 风类要素和温压类要素各自之间相关性较高, 其他要素间相关性不明显。提取各要素与跑道视程的相关性(不考虑正负相关)从大到小排列, 结合专家经验, 本文选取 14 个相关性较强要素作为预报输入因子, 依次分别为 2 分钟平均风速、2 分钟平均风向、10 分钟平均风速、10 分钟平均风向、背景亮度、10 分钟平均气象光学视程、1 分钟平均气象光学视程、10 分钟平均跑道视程、1 分钟平均跑道视程、1 分钟平均降水量、露点温度、气压、相对湿度和温度。数据格式如表1所示。

以上输入预报因子的选取, 不仅确保了对现有观测资料的充分使用, 也能体现各要素之间的相互影响作用随时间的动态变化过程, 进而构造一个比较完备的预测数据集。

## 2.3 多变量时序预测

多变量时序预测问题 (multivariable time series prediction problem) 是指用过去的多个变量的观测值预测其中某一个或多个变量未来的观测值。多步预测 (multi-step forecasting) 是指用过去的观测值预测未来一个时间序列的观测值<sup>[18]</sup>。

本文利用预报时间点前 4 个小时的本场 AWOS 自动观测的 14 个预报因子作为输入, 预测未来 6 小时每分钟的 RVR 值, 即多步多变量时序预测。

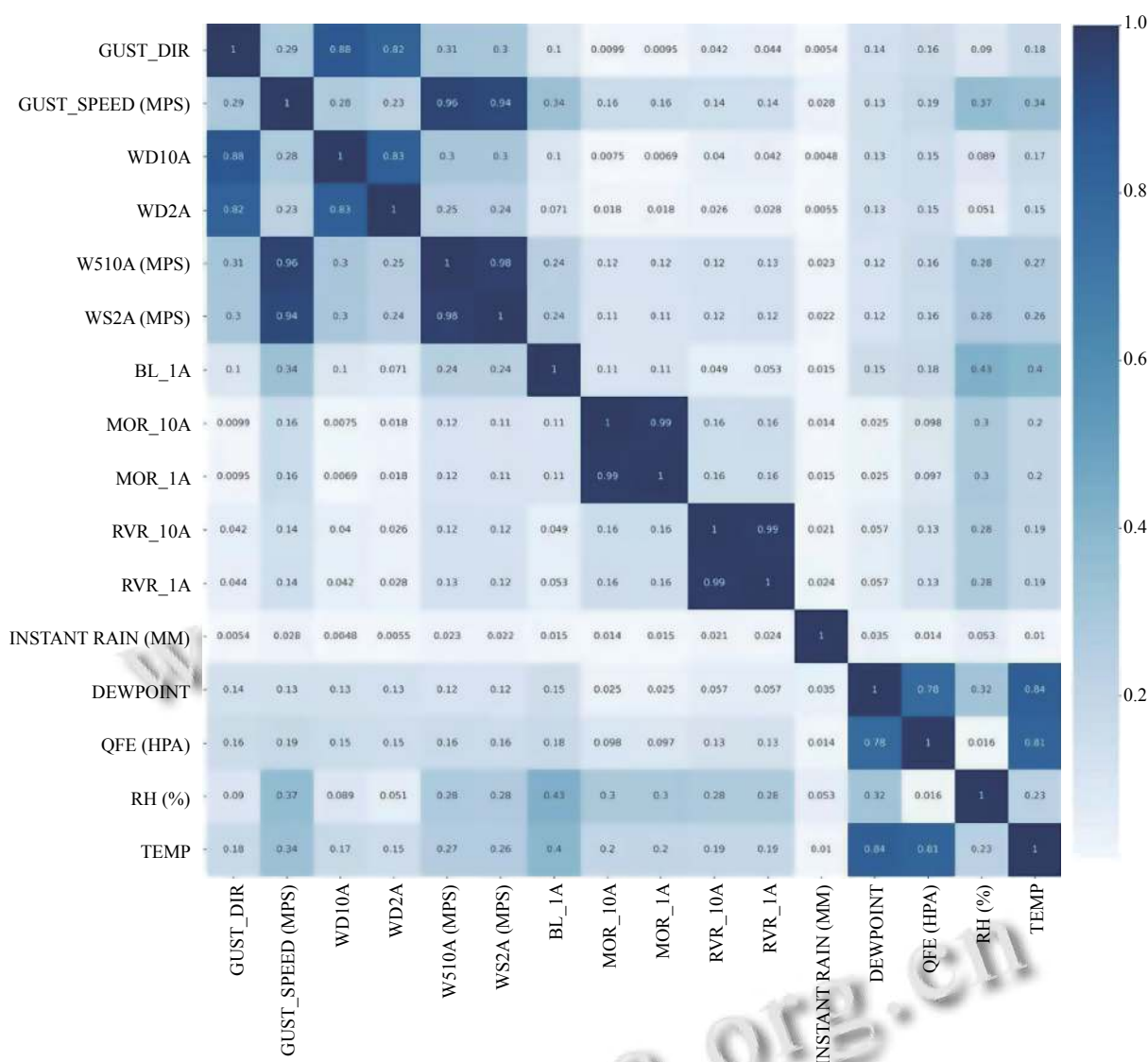


图4 AWOS地面观测要素之间的相关性图

表1 数据格式

时间	因子1	因子2	...	因子14
2012-1-1 0:00	341	0.12	...	87
...	...	...	...	...
2018-12-31 23:59	320	0.25	...	87

时间序列数据集是由按时间索引排序的一系列数字组成,为了能够从输入模式里学习如何预测输出模式,需要将时间序列数据集转化为监督学习的数据格式,即监督学习数据集由输入模式X和输出模式Y构成.将跑道视程预测时间序列数据集转化为监督学习的数据格式,如表2所示.

如表2所示,这是多步多变量时序预测的监督学习数据格式,表2中t表示预报时间点,t-240表示预报

时间点过去240分钟,即过去4个小时,t+359表示预报时间未来359分钟,即未来6个小时.表2左半部分即监督学习数据集中的输入模式X,其中包括过去4个小时的14个预报因子观测值;表2右半部分即监督学习数据集中的输出模式Y,未来6个小时的RVR观测值.

将跑道视程预测的时间序列数据集转化为监督学习数据集之后,对数据集进行数据预处理,包括缺省值处理、归一化、样本均衡等.

由于各输入的预报因子存在量级和单位上的差异,为了消除各个预报因子之间的差异带来的数据不均一性,需要对跑道视程预测数据集作标准的归一化处理,使每一个预报因子的数据落在区间[-1,1]内.

由于低跑道视程天气在整个数据集中所占比例较小, 样本均衡就格外重要, 本文采用了分层抽样法进行样本均衡. 分层抽样法就是将总体单位按其属性特征分成若干类型或层, 然后在类型或层中随机抽取样本

单位. 分层抽样的特点是由于通过划类分层, 增大了各类型中单位间的共同性, 容易抽出具有代表性的调查样本<sup>[19]</sup>. 该方法适用于总体情况复杂, 各单位之间差异较大, 单位较多的情况.

表2 监督学习数据格式

时间	因子1 ( $t-240$ )	因子2 ( $t-240$ )	...	因子14 ( $t-240$ )	...	因子1 ( $t-1$ )	...	因子14 ( $t-1$ )	$RVR(t)$	...	$RVR(t+359)$
2012-1-1 3:59	341	0.12	...	87	...	359	...	89	1100	...	1000
...	...	...	...	...	...	...	...	...	...	...	...
2018-12-31 17:59	320	0.25	...	87	...	298	...	90	1000	...	800

将跑道视程预测数据集输出模式 $Y$ 中的 $t$ 时刻 $RVR$ 值, 按照表3  $RVR$ 等级进行划分<sup>[20]</sup>, 分成5层, 然后在这5层中进行随机抽样, 从而使跑道视程预测数据集中的样本均衡.

表3  $RVR$ 等级

等级	$RVR$ 等级范围 (m)
0	$RVR \leq 200$
1	$200 < RVR \leq 350$
2	$350 < RVR \leq 500$
3	$500 < RVR \leq 800$
4	$RVR > 800$

将经过数据预处理(缺省值处理、归一化、样本均衡)的跑道视程预测数据集按照6:2:2的比例进行随机抽样, 并将样本分组为训练集、验证集和测试集. 训练集被用于预测模型构建, 验证集被用于模型参数最优调试和测试集被用于模型预测效果检验, 从而保障模型的可信性和科学性.

### 3 机场跑道视程预测实证分析

#### 3.1 LSTM 预测模型构建

机场跑道视程预测问题为多步多变量时序预测问题, 因此本文采用了适合处理时序数据的长短期记忆网络, 提出了一种基于LSTM的机场跑道视距预测模型.

根据柯尔莫哥洛夫理论(Kolmogorov theory), 当前馈型神经网络具有3层及以上结构时可实现以任意精度逼近任意非线性连续函数<sup>[13,21,22]</sup>. 因此, 本文的预测模型由3个LSTM层和1个全连接的神经网络层组成, 且为了避免过拟合, 在每个LSTM层后面添加了dropout层, 结构图如图5所示. 过去4个小时的气象要素(风速、温度、湿度和跑道视程等)被输入网络, 预测跑道未来6个小时的视程.

图5中 $X_{t-i}$ 表示 $t-i$ 时刻14个预报因子的观测值,

$X_{t-i} = [x_1^{t-i}, x_2^{t-i}, \dots, x_{14}^{t-i}]$ ,  $i \in [1, 240]$ , 因此输入样本维度为 $[None, 240, 14]$ ,  $None$ 表示数据集样本数量;  $H_{t-i}$ 表示 $t-i$ 时刻LSTM cell的输出值,  $H_{t-i} = [h_1^{t-i}, h_2^{t-i}, \dots, h_n^{t-i}]$ ,  $i \in [1, 240]$ , 因此LSTM层的数据输出维度为 $[None, 240, n]$ , 其中,  $n$ 表示LSTM基本节点中Sigmoid层与tanh层的输出维度, 也就是LSTM cell的输出维度, 需要人为设定, 下面将针对LSTM cell的实际输出维度进行探讨. Flatten表示压缩, 将LSTM层的输出数据从三维 $[None, 240, n]$ 压缩至二维 $[None, 240 \times n]$ , 然后将 $[None, 240 \times n]$ 的二维数据送入dense全连接层, dense层包括360个神经元, 从而输出未来6个小时每分钟的 $RVR$ 值, 因此dense层的数据输出 $Y$ 维度为 $[None, 360]$ .

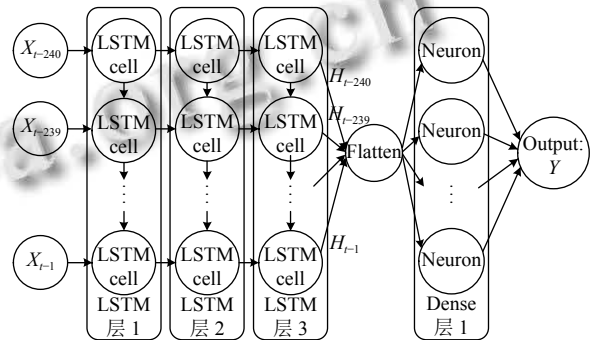


图5 跑道视程预测模型网络结构图

LSTM的遗忘门、输入门和输出门都是通过图5中LSTM cell实现, LSTM cell具体算法流程如下:

输入: 上一个状态的信息, 假设输入门表示为 $i_t$ , 遗忘门表示为 $f_t$ 和输出门表示为 $o_t$ , 下标 $t$ 表示时刻, 使用 $C_t$ 表示LSTM cell在 $t$ 时刻的单元状态<sup>[15]</sup>;

步骤1. 接收输入信息. 通过当前 $t$ 时刻输入数据 $X_t$ 、上一时刻LSTM cell输出值 $H_{t-1}$ 与上一时刻单元状态 $C_{t-1}$ , 计算输入门的值 $i_t$ , 输入门决定了将多少当前

时刻输入数据 $X_t$ 的信息保存到单元状态 $C_t$ ;

步骤 2. 遗忘状态信息. 遗忘门用于控制上一个时刻单元状态 $C_{t-1}$ 的哪些信息保留到当前时刻的单元状态 $C_t$ 中;

步骤 3. 更新状态信息. 通过输入门决定当前时刻单元状态哪些信息需要被保留, 结合遗忘门新的需要添加的信息, 对当前时刻的单元状态进行更新, 得到新的当前时刻单元状态 $C_t$ ;

步骤 4. 输出新的信息. 输出门主要是通过 Sigmoid 层决定要输出细胞状态的相关信息, 然后将输出门的输出 $o_t$ 与  $\tanh$  处理后的记忆单元状态 $C_t$ 进行乘积, 得到最终当前层的输出信息 $H_t$ ;

输出: 当前 LSTM 单元的状态信息 $C_t$ 及输出 $H_t$ .

预测模型网络结构确定之后, 为充分发挥 LSTM 网络的非线性映射能力, 将它们的 cell 输出维度 $n$ 分别从 1 递增到 1 000, 并逐个进行实验, 寻找最适合本研究的 LSTM cell 输出维度组合. 基于 LSTM 的跑道视程预测模型的构建过程如图 6 所示.

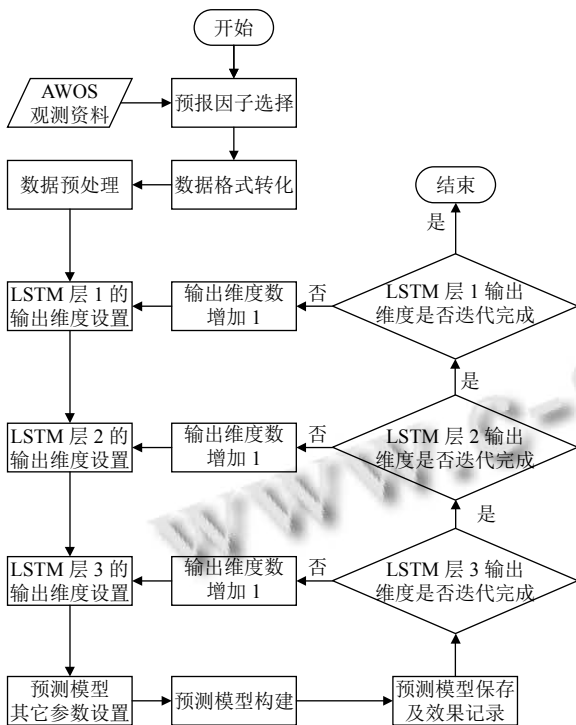


图 6 基于 LSTM 的跑道视程预测模型构建流程图

为了避免随机初始权重和偏置对预测结果的影响, 分别将每种 LSTM 输出维度组合对应的预测模型进行 100 次训练和测试, 保存其效果最好时的网络参数.

通过实验发现, 随着 LSTM 层输出维度的增加, 预

测模型效果先急剧上升, 然后渐趋平缓, 特别地, 当 LSTM 层 1 的输出维度大于 490、LSTM 层 2 的输出维度大于 420、LSTM 层 3 的输出维度大于 360 时, 继续增加输出维度给预测模型带来的效果提升几乎可以忽略不计. 综合考虑计算代价增量与效果提升幅度后认为: LSTM 层 1、LSTM 层 2 与 LSTM 层 3 的输出维度数分别取 490、420 与 360 时最适合本研究.

其他相关的实验参数设置如下: 随机初始化权重和偏置, 损失函数为训练集的均方误差, 损失函数的优化算法采用 RMSProp 优化器, dropout 比率为 0.2, 终止条件为训练集迭代次数达到 1 万次或者均方误差小于  $1E-5$ .

该网络采用基于 TensorFlow 的 Keras API 实现, 并采用 4 块 NVIDIA TiTAN V GPU 进行训练.

### 3.2 预测试验结果分析

AWOS 观测设备分别位于两条跑道的 4 个跑道端点及 2 个中点, 2 个跑道中点的 AWOS 观测设备只对少量气象要素进行监测, 缺少与 RVR 相关性较强的预报因子 (降水量、气压、相对湿度、温度等), 因此无法进行跑道视程预测.

将两条跑道 4 个跑道端点 RWY23L、RWY05R、RWY23R、RWY05 的 AWOS 观测数据经过数据预处理 (缺省值处理、归一化、样本均衡) 后, 按照 6:2:2 比例进行随机抽样, 划分成训练集、验证集和测试集, 送入第 3.1 节构建的 LSTM 跑道视程预测模型进行训练, 4 个跑道端点预测模型的训练及验证 Loss 曲线如图 7.

通过图 7 可以看出, 随着迭代次数的增加, 4 个跑道端点的 Loss 值都在下降, 训练集下降比较平滑, 最后趋于稳定, 验证集下降过程起伏较大.

将测试集送入训练后的预测模型, 对 4 个跑道端点的跑道视程实际观测值与预测值进行比较, 分别计算拟合度及均方根误差. 拟合度 (goodness of fit) 是指预测值对观测值的拟合程度<sup>[23]</sup>, 一般使用变量  $R^2$  表示,  $R^2$  值越接近 1, 说明拟合程度越好, 拟合度  $R^2$  计算方法: 设  $y$  为待拟合值, 其均值为  $\bar{y}$ , 拟合值为  $\hat{y}$ , 记总平方和 SST:

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (9)$$

回归平方和 SSR:

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \quad (10)$$

残差平方和SSE:

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (11)$$

则 $SST = SSR + SSE$ , 拟合度 $R^2$ :

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \quad (12)$$

均方根误差 (root mean square error,  $RMSE$ ) 是反映估计量与被估计量之间差异程度的一种度量<sup>[24]</sup>, 均方根误差越小越好. 均方根误差 $RMSE$ :

$$RMSE = \sqrt{E(\hat{y} - y)^2} = \sqrt{\frac{SSE}{n}} \quad (13)$$

未来 6 小时 0-359 分钟的预测模型拟合度 $R^2$ 及均方根误差 $RMSE$ 变化过程如图 8 所示.

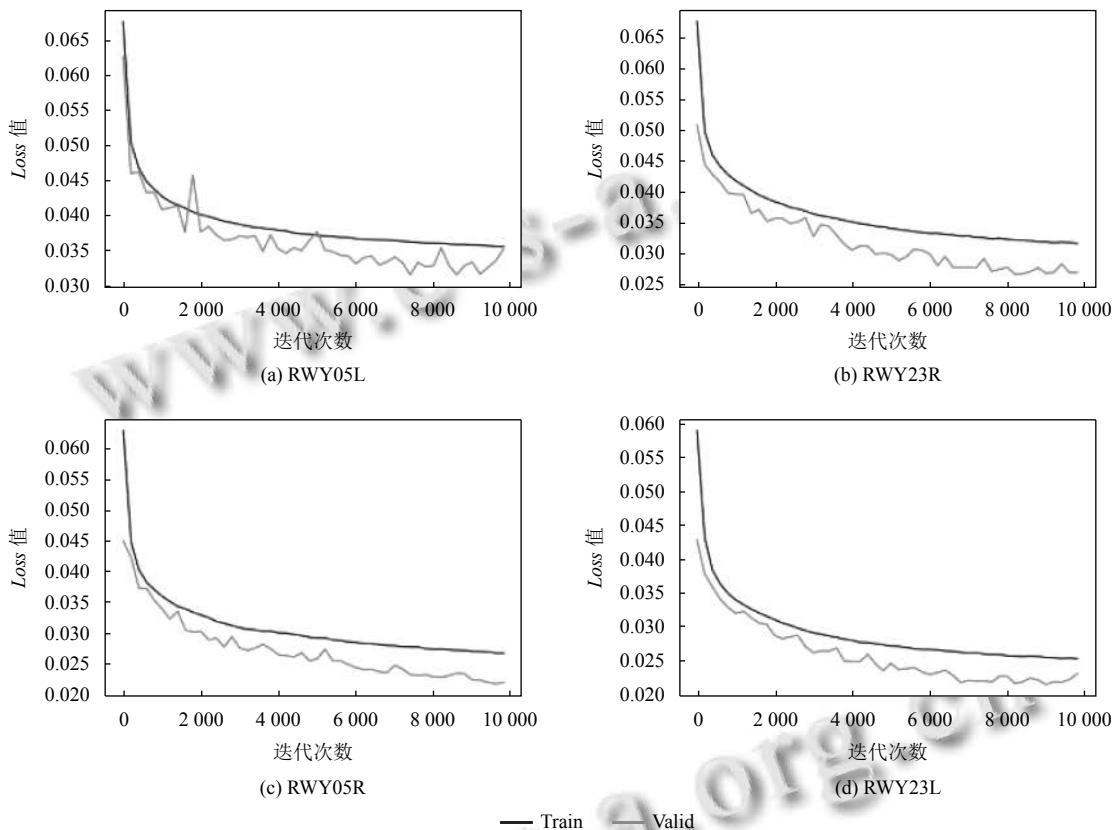


图 7 4 个端点训练集及验证集的Loss曲线

图 8 可见, 跑道视程预测模型在 0-2 小时内平均拟合度能够达到 72%, 随着时间的增加, 预测模型的拟合度 $R^2$ 逐渐减小, 最后降至 50% 以下, 并且均方根误差 $RMSE$ 逐渐增大, 最后增至 438.471 以上, 这说明跑道预测模型预测较近时间点的 $RVR$ 值具有较高的拟合度, 较低的均方误差, 预测得比较准确, 而预测未来较远时间点的 $RVR$ 值偏差会越来越大, 这也符合直观认知, 跑道视程预测比较适合 0-2 小时临近预报.

#### 4 小结与讨论

跑道视程的时序分布特征等分析结果表明影响咸阳机场低跑道视程的天气主要是大雾. 由于大雾的局地

特性很强, 其预报目前在气象领域依然是业务难点. 虽然近年来大雾预报由传统的经验或者统计预报向数值模式或者大雾模式预报转变, 但数值释用和雾模式预报准确性仍有待提高, 难以在实际业务预报中直接使用<sup>[25]</sup>.

考虑到每个预测时间前 4 小时的多维气象要素数据包含着各种气象要素的变化趋势, 例如降水增大, 未来跑道视程数据就会相应的降低, 湿度增大、温度降低、风速较小, 一般会伴随着大雾的发生 (大雾的机理比较复杂, 湿度增大、温度降低、风速较小只是其中 3 个较为显著的气象条件), 未来跑道视程也会相应的降低. 本文基于长短时记忆网络的多维时间序列预测方法, 采用多维气象要素进行跑道视程预测, 旨在拟合



出多维气象要素与跑道视程之间的映射关系。

实验表明采用长短时记忆网络,利用风速、温度、湿度和跑道视距等预报因子建立一种机场跑道视程预测模型,对测试集预测的0-2小时拟合度在58%—

88%之间,拟合度结果表明本文搭建的LSTM预测模型对跑道视程具有较高的预测能力,具有推广价值和前景,其结果可被航空气象保障人员参考应用LSTM的优越性和推广前景。

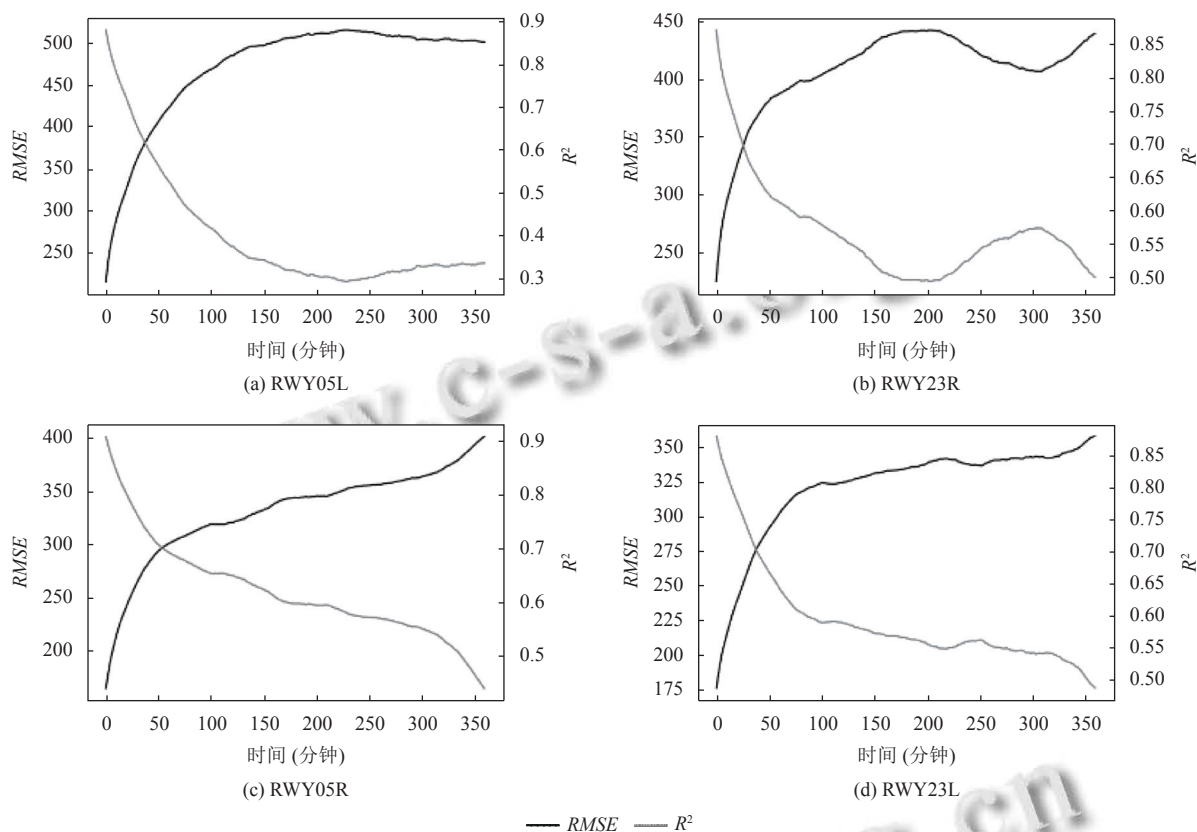


图8 4个端点未来0-359分钟模型 $R^2$ 及RMSE变化图

未来研究中将要改进的方面:

(1) 低跑道视程天气的出现与季节、晨昏时间信息有关,之后的研究工作中预报因子的选取考虑季节、晨昏等时间信息。

(2) 冬季出现较低跑道视程时,边界层内常出现大气层结中性弱不稳定、逆温以及大湿度区分布的垂直结构.为了尽可能地描述大气垂直结构对低跑道视程天气的影响,将增加探空观测的多层次的位势高度、温度、露点温度、风向、风速等资料作为预报因子<sup>[25]</sup>。

(3) 随着人工智能技术的发展,越来越多优秀的预测技术逐渐成熟,采用最新的预测算法建模,进行比较,选取较优的预测模型为跑道视程的预测提供指导,更好的保障航空器的安全起降。

(4) 虽然0-2小时跑道视距预测具有很高的拟合

度.然而,预测模型有几个部分任然需要进行改进,如跟踪谷值<sup>[26]</sup>,跑道视程谷值是航空管制员重点关注的,需要根据跑道视程谷值判断飞机是否能够起降。

#### 参考文献

- 1 中国民用航空总局.中国民用航空仪表着陆系统II类运行规定. [http://www.law-lib.com/law/law\\_view.asp?id=12814](http://www.law-lib.com/law/law_view.asp?id=12814). (1996-10-16).
- 2 杨晶轶.成都双流机场低跑道视程特征及其与气象要素的关系.高原山地气象研究,2018,38(2):90-95. [doi: 10.3969/j.issn.1674-2184.2018.02.015]
- 3 朱蕾,朱国栋,石江.SVM方法在乌鲁木齐机场跑道视程预测中的应用.2007年中国气象学会年会论文集.广州:中国气象学会,2007.17-25.
- 4 Ku SK, Park CH, Hong SM. Advanced estimation model of

- runway visual range using deep neural network. *The Journal of Advanced Navigation Technology*, 2018, 22(6): 491–499.
- 5 周志洋. 基于数据挖掘的高高原机场跑道视程短时预测方法研究 [硕士学位论文]. 成都: 电子科技大学, 2018.
  - 6 袁芳. 人工智能在气象领域的应用与建议. *科学咨询/科技管理*, 2021, (10): 84–85.
  - 7 李芬, 杨程, 赵晋斌, 等. 基于天气类型聚类和 LSTM 的  $PM_{2.5}$  短期预测模型. *水电能源科学*, 2021, 39(3): 199–202, 151.
  - 8 魏昱洲, 许西宁. 基于 LSTM 长短期记忆网络的超短期风速预测. *电子测量与仪器学报*, 2019, 33(2): 64–71.
  - 9 罗芳琼, 陆虹, 秦川, 等. 基于 LSTM NN 特征提取的低温气象灾害预报方法研究. *数学的实践与认识*, 2021, 51(10): 169–182.
  - 10 中国民用航空局. 民用航空气象地面观测规范. [http://www.caac.gov.cn/XXGK/XXGK/GFXWJ/201511/t20151102\\_8169.html](http://www.caac.gov.cn/XXGK/XXGK/GFXWJ/201511/t20151102_8169.html). (2012-02-28).
  - 11 中国民用航空总局. 航空器机场运行最低标准的制定与实施规定. [http://www.gov.cn/gongbao/content/2002/content\\_61982.htm](http://www.gov.cn/gongbao/content/2002/content_61982.htm). (2001-02-26).
  - 12 Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*. 1997, 9(8): 1735–1780.
  - 13 李文静, 王潇潇. 基于简化型 LSTM 神经网络的时间序列预测方法. *北京工业大学学报*, 2021, 47(5): 480–488.
  - 14 Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, 9(8): 1735–1780.
  - 15 於雯, 周武能. 基于 LSTM 的商品评论情感分析. *计算机系统应用*, 2018, 27(8): 159–163. [doi: 10.15888/j.cnki.csa.006483]
  - 16 余杰杰. 基于 LSTM 的道路拥堵分析模型研究 [硕士学位论文]. 南昌: 华东交通大学, 2019.
  - 17 Shi XJ, Chen Z, Wang H, *et al.* Convolutional LSTM network: A machine learning approach for precipitation nowcasting. arXiv: 1506.04214, 2015.
  - 18 Liu YW, Gong CY, Yang L, *et al.* DSTP-RNN: A dual-stage two-phase attention-based recurrent neural network for long-term and multivariate time series prediction. *Expert Systems with Applications*, 2020, 143: 113082.
  - 19 Imbens GW, Lancaster T. Efficient estimation and stratified sampling. *Journal of Econometrics*, 1996, 74(2): 289–318.
  - 20 李跃春, 周伦波. 双流机场 02 号 RVR 分析和 II 类运行气象预报服务探讨. *四川气象*, 2007, 27(S1): 38–42.
  - 21 Kůrková V. Kolmogorov's theorem and multilayer neural networks. *Neural Networks*, 1992, 5(3): 501–506.
  - 22 Hecht-Nielsen R. Kolmogorov's mapping neural network existence theorem. *Proceedings of the 1st IEEE International Conference on Neural Networks*. San Diego: IEEE, 1987, 3: 11–14.
  - 23 Mcelroy MB. Goodness of fit for seemingly unrelated regressions: Glahn's  $R_{yx}^2$  and Hooper's  $r^2$ . *Journal of Econometrics*, 1977, 6(3): 381–387.
  - 24 Rudoy GI. On modification of the mean squared error loss function for solving nonlinear heteroscedastic errors-in-variables problems. *Informatics and Applications*, 2017, 11(2): 74–84.
  - 25 朱蕾, 朱国栋. 支持向量机方法在机场跑道视程预报中的应用. *暴雨灾害*, 2010, 29(2): 171–175.
  - 26 冯汉中, 陈永义, 成永勤, 等. 双流机场低能见度天气预报方法研究. *应用气象学报*, 2006, 17(1): 94–99. [doi: 10.3969/j.issn.1001-7313.2006.01.013]