

兰明才,唐杰,周莉,等.基于逐步回归模型的华南区域模式气温预报产品释用技术研究[J].气象与环境科学,2023,46(5):86-95.  
Lan Mingcai,Tang Jie,Zhou Li,et al.Research on Interpretation and Application of Temperature Forecast Products of Regional Model in South China Based on Stepwise Regression Model[J].Meteorological and Environmental Sciences,2023,46(5):86-95.  
Doi: 10.16765/j.cnki.1673-7148.2023.05.011

## 基于逐步回归模型的华南区域模式气温 预报产品释用技术研究

兰明才<sup>1,2</sup>,唐杰<sup>1,2</sup>,周莉<sup>1,2</sup>,陈静静<sup>1,2</sup>,谭萍<sup>3</sup>,刘青竹<sup>4</sup>

(1.湖南省气象台,长沙 410118; 2.气象防灾减灾湖南省重点实验室,长沙 410118;  
3.湖南省气象服务中心,长沙 410118; 4.集美区气象局,厦门 361021)

**摘 要:** 基于2017–2019年华南高分辨率区域数值模式2 m 气温预报产品,设计了3套模型的构建方案,利用25个显著因子建立97个国家基本站的逐步回归订正模型,对模式和3套订正模型的预报与实况进行偏差分析,结果表明:(1)华南区域模式分辨率较高,但无法细致反映湖南复杂地形影响导致的区域气温变化特征,模式预报的气温,山区的偏差大于平原的;在1–72 h 预报时效内的准确率为54.08%,预报偏差随预报时效延长显著增加,且具有明显的周期振荡,偏差最大值出现在日最高/最低气温发生时段。(2)基于逐步回归模型,依据模式的起报时次、预报时效建立了3套气温订正模型。其中,方案1为不分起报时次和预报时效建模,方案2为分起报时次不分预报时效建模,方案3为分起报时次和预报时效建模。3套针对不同起报时次和预报时效的建模方案皆对湖南2 m 气温的时空分布特征有较好刻画,对复杂地形影响区域的气温预报有一定的订正能力,但随着预报时效的延长订正效果下降,订正效果总体为00时起报的优于12时起报的。(3)3套建模方案2 m 气温预报同样具有周期振荡特征,但没有明显的季节差异,最大偏差出现在最高气温发生时段,对夜间温度,尤其是对05时前后的温度预报订正效果明显,其中方案1的气温订正效果总体优于方案2和方案3的,1–72 h 预报时效内绝对误差减小了0.68℃,准确率提高了17%以上。

**关键词:** 逐步回归; 华南区域模式; 2 m 气温; 气温订正

中图分类号: P459.9

文献标志码: A

文章编号: 1673-7148(2023)05-0086-10

### 引 言

随着社会发展和人民生活水平的提高,对气温预报的精准度提出新的要求,气温的准确预报对热浪、霜冻等灾害性天气的防灾减灾起重要的指导作用<sup>[1-3]</sup>。湖南地形呈马蹄状,呈平原与三面环山的地形特点,受复杂的天气系统影响,精细化温度预报成为湖南区域天气预报的难点之一。随着数值预报和资料同化技术的发展,数值预报准确率越来越高,在天气预报业务中发挥着越来越重要的作用<sup>[4-7]</sup>。

由于模式自身的系统误差及特殊地形和下垫面等的影响,需要对数值预报产品进行订正释用,以提高模式气温预报的准确率<sup>[8-12]</sup>。

对于模式气温预报偏差的原因分析<sup>[13-16]</sup>一直被学术界和数值模式应用部门热衷探讨,并围绕数值预报产品释用技术开展了大量研究,包括MOS、BP神经网络、卡尔曼滤波、递减平均法等多种技术已取得显著进展并得到较好应用<sup>[17-22]</sup>。随着观测数据和模式产品种类的增多,常规方法难以应对数据多样性的弊端逐渐显现,近年来机器学习方法在气象领域得到迅速应用。陈昱文等<sup>[3]</sup>提出一种基

收稿日期: 2020-11-25; 修订日期: 2023-04-13

基金项目: 湖南省自然科学基金项目(2022JJ40214); 2023 中国气象局复盘专项(FPZJ2023-088); 湖南省气象局第三期业务能力建设项目(NLJS01、NLJS02) 共同资助

作者简介: 兰明才(1984-),男,湖南株洲人,高级工程师,硕士,从事天气预报预测及短时临近预报预警技术研究.E-mail: 372917798@qq.com

通信作者: 唐杰(1980-),男,湖南永州人,正研级高工,硕士,从事天气气候诊断研究工作.E-mail: 7101768@qq.com

于数值模式预报产品的气温预报集成学习误差订正方法,通过人工神经网络、长短期记忆网络和线性回归模型组合出新的集成学习(ALS)模型,有效地提升了站点气温预报整体准确率。这些技术各具特点,但订正效果仍受到相关条件的限制,使其具有一定的适用性,且以上研究大多基于ECMWF、GRAPES和GFS等全球大尺度预报模式,各区域中心、发达省市发展的中尺度预报模式适用技术相对较少。

伴随互联网云技术的发展,湖南在泛珠江三角洲区域模式预报发展联盟框架下,实现了华南区域模式的图形和数据快速共享。华南区域模式具有更新快、精度高等特点。李刚等<sup>[23]</sup>研究发现,华南3 km分辨率区域模式对贵州气温预报具有较好的参考价值,最高(低)气温及35.0℃以上高温事件预报准确率均在80%左右。但华南3 km区域模式在其他地区尤其是湖南地区的预报评估鲜有提及。为更好地应用该模式,有必要对其在湖南的预报偏差进行分析,了解华南区域模式在湖南的适用性,并有针对性地发展基于该区域模式的释用技术,提升华南区域模式在湖南的精细化温度预报能力。

回归分析技术是确定两种或两种以上变量间相互依赖的定量关系的一种统计分析方法,由于其受到相关条件的限制较少,在气温订正中运用十分广泛。康岚等<sup>[24]</sup>利用GRAPES中尺度模式对西南地区2 m气温预报进行检验,发现回归分析技术对预报模式订正在一定程度上能改进预报效果。李佰平等<sup>[25]</sup>采用一元线性回归、多元线性回归法对模式地面气温预报进行订正,发现能有效地减小地面气温多个时效预报的误差。王敏等<sup>[26]</sup>采用非齐次高斯回归技术,对2 m气温预报结果开展了一阶偏差和二阶离散度的校准研究,校准后的2 m气温预报可靠性和预报技巧均显著提高。李超等<sup>[27]</sup>基于ECMWF模式温度预报和浙江省72个国家基本站观测资料采用K-近邻(KNN)回归算法改进温度预报,发现KNN方案对地形复杂地区的温度预报改进效果明显,对冷空气活动和夏季高温等天气过程预报改善效果也较稳定。但以上研究针对预报因子的选择基本依赖于预报经验,尤其是针对华南区域模式物理量相对较多且研究不足,如何保证在已选定的因子中得到最优的回归方程是难点。殷悦等<sup>[28]</sup>研究发现相较于传统单元线性回归,利用多因子的逐步回归能有效提高延长数据的准确性。本文在前人的研究基础上,利用逐步回归分析方法解决这一

难题,针对华南区域模式(新版MARS3 km)2 m气温预报产品,利用25个显著因子建立基于站点的逐步回归订正模型,对订正前后的模式预报与实况进行偏差分析,为预报员智能精细化网格温度预报提供客观参考依据。

## 1 资料与方法

### 1.1 数据

本文所用资料包括湖南97个国家站逐1 h 2 m气温实况,华南中尺度天气预报模式(新版MARS3 km)<sup>[29-32]</sup>每日00/12 UTC(世界时,北京时为08/20时)输出的逐1 h 2 m气温预报产品,最优距离反比插值到湖南97个站点。模式预报与实况数据选取时段为2017年3月至2019年12月,模式起报时间为00时、12时(世界时),预报时效为起报时间至84 h,2017-2018年数据作为训练集,2019年数据为检验数据集。根据业务实际需求,建模预报时间为起报时间往后的13-84 h预报(为方便对结果进行分析,下文以1-72 h代称)。

### 1.2 逐步回归法

采用逐步回归法引入多个变量,引入后进行检验,删除不显著变量,最终得到每个变量均显著的逐步回归方程。逐步回归方程如下:

$$\begin{cases} y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_p x_{1p} + e_1 \\ y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_p x_{2p} + e_2 \\ \vdots \\ y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_p x_{np} + e_n \end{cases} \quad (1)$$

其中 $\beta_0, \beta_1, \cdots, \beta_n$ 为 $p+1$ 个待估计参数, $x_0, x_1, \cdots, x_n$ 是 $p$ 个一般变量,设 $e_0, e_1, \cdots, e_n$ 是 $n$ 个相互独立的且遵从同一正态分布 $N(0, \sigma)$ 的随机变量,为简明起见,这一模型写成矩阵形式:

$$y = X\beta + e \quad (2)$$

其中 $y, \beta, e$ 为向量,分别为

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}, e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} \quad (3)$$

基于华南中尺度天气预报模式输出的地面和高空要素,以及基于要素计算出常用的温度平流、散度、涡度等物理量(共165个因子),利用逐步回归方法在湖南97个国家站上分别建立温度预报模型。考虑到业务应用的便利性,选取影响站点数超过20%(即19站次)的25个因子引入逐步回归方程,初选因子见表1。该25个因子在97个站点模型中

出现的频次都超过 19 站次,其中应用频次最高的为 2 m 气温,在 97 个站点模型中都得到应用,应用频次最低的为 850 hPa 经向风和 925 hPa 散度,在 21 个站点模型中都得到应用。利用 25 个因子在各个

站点建立温度预报模型,每个站点模型最终用到的因子和因子数由逐步回归法自动拟合。剔除不显著因子,大部分站点建立的逐步回归方程最终用到的显著因子数平均为 6~10 个。

表 1 逐步回归法挑选的显著因子

地面	总降水( 27 )、10 m $u$ 风( 73 )、10 m $v$ 风( 50 )、2 m 相对湿度( 41 )、2 m 气温( 97 )
高空	纬向风 $u$ ( 700 hPa( 48 ) /500 hPa( 50 ) )
	经向风 $v$ ( 925 hPa( 22 ) /850 hPa( 21 ) /500 hPa( 28 ) )
	温度 $T$ ( 925 hPa( 22 ) /850 hPa( 22 ) /700 hPa( 33 ) /500 hPa( 54 ) )
	相对湿度 RH( 1000 hPa( 45 ) /925 hPa( 28 ) /850 hPa( 37 ) /600 hPa( 44 ) /500 hPa( 30 ) )
	温度平流( 1000 hPa( 31 ) /850 hPa( 38 ) /700 hPa( 23 ) )
	散度( 1000 hPa( 64 ) /925 hPa( 21 ) )
	涡度( 700 hPa( 34 ) )

注: 括弧内的数字表示该因子在 97 个站点模型中的应用频次,单位: 站次。

### 1.3 建模方案

在日常业务中不难发现,不同起报时次存在不一致的系统性偏差。王丹等<sup>[32]</sup>研究同样发现 GRAPES 模式 08 时起报的准确率高于 20 时起报

的。此外,大部分研究都表明,随着预报时效的增加,预报准确率差异较大<sup>[19-21, 24-30]</sup>。因此,本文设计了 3 套模型的构建方案(表 2),以分析不同起报时次和预报时效差异对模型预报效果的影响。

表 2 建模方案

方案名称	方案设计路线
方案 1	不考虑起报时次和预报时效差异,00/12 时起报的 1-72 h 预报构建 1 个统一模型
方案 2	不考虑预报时效差异,按照 00/12 时 2 个起报时次分别构建 2 个模型
方案 3	充分考虑不同起报时次和预报时效差异,按照 00/12 时 2 个起报时次,以及 1-72 h 逐 1 h 预报时效分别构建 144 个模型

### 1.4 检验方法

以相对误差和绝对误差讨论模式预报订正结果、模式预报与实况的偏离程度。以订正或预报结果与实况温度差值的绝对值 $\leq 2$ 作为预报正确的标准,计算预报准确率,探讨逐步回归模型对 2 m 温度的订正效果。

## 2 订正结果分析

### 2.1 3 套建模方案的气温订正整体效果对比

表 3 为不同建模方案的 2 m 气温订正在 4 个预报时效(1-72 h、1-24 h、25-48 h、49-72 h)的准确率。除了方案 2 中 49-72 h 预报时效及方案 3 中 25-48 h 时效外,其余预报时效内 3 套方案的订正效果都

是 00 时起报的 2 m 气温预报准确率明显优于 12 时起报的。对于 00 时起报的温度预报,方案 1 在 4 个预报时效的订正效果都是最佳的,其次是方案 2,方案 3 的订正效果最差。对于 12 时起报的温度预报,方案 1 在 1-72、1-24 h 及 49-72 h 预报时效的订正效果最佳,分别达到 70.49%、74.59% 和 65.94%,方案 2 在 25-48 h 预报时效的订正效果最佳,为 69.12%,方案 3 在所有预报时效内订正效果都是最差的。此外,无论是 00 时起报还是 12 时起报,3 套建模方案气温订正准确率随着预报时效的延长是下降的,1-24 h 的预报准确率最高,预报质量明显优于对 25-48 h 和 49-72 h 的预报。

表 3 不同建模方案的 2 m 气温订正准确率

%

预报时效	方案 1		方案 2		方案 3	
	00 时起报结果	12 时起报结果	00 时起报结果	12 时起报结果	00 时起报结果	12 时起报结果
1-72	71.22	70.49	69.92	70.05	67.31	66.98
1-24	75.96	74.59	74.88	74.03	73.98	73.94
25-48	71.38	69.01	69.52	69.12	64.93	65.02
49-72	66.01	65.94	65.80	65.91	62.87	61.74

总体而言,方案 1 的气温订正预报准确率高于方案 2 和方案 3 的。以下分析方案 1 较原模式的订

正效果。表 4 为 00 时次起报的模式预报及逐步回归订正后预报的 2 m 气温对比。72 h 内模式预报的

温度总体偏低  $0.73^{\circ}\text{C}$  ,绝对误差为  $2.25^{\circ}\text{C}$  ,  $2^{\circ}\text{C}$  内的准确率为  $54.08\%$  。通过逐步回归订正,修正了模式预报总体偏低的情况,较模式预报的绝对误差减小了  $0.68^{\circ}\text{C}$  ,准确率提高了  $17.14\%$  。从逐  $3\text{ d}$  的预报检验看,经过逐步回归模型订正后,逐  $3\text{ d}$  的温度

预报效果都有不同程度的提高,尤其是  $25\sim 48\text{ h}$  内的订正效果最显著,较原模式预报的相对误差和绝对误差分别减少了  $0.44$  和  $0.66$  ,准确率提高了  $18.92\%$  ,对  $49\sim 72\text{ h}$  预报的订正效果是最差的,但准确率也提高了  $14.02\%$  。

表4 方案1逐步回归订正前后温度预报检验结果(00时次起报)

时效	逐步回归			原模式			订正效果		
	相对误差 / $^{\circ}\text{C}$	绝对误差 / $^{\circ}\text{C}$	准确率 / $\%$	相对误差 / $^{\circ}\text{C}$	绝对误差 / $^{\circ}\text{C}$	准确率 / $\%$	相对误差 / $^{\circ}\text{C}$	绝对误差 / $^{\circ}\text{C}$	准确率 / $\%$
1-72 h	0.03	1.57	71.22	-0.73	2.25	54.08	0.70	0.68	17.14
1-24 h	0.05	1.34	75.96	-0.78	2.01	57.56	0.83	0.67	18.40
25-48 h	-0.01	1.67	71.38	-0.45	2.33	52.46	0.44	0.66	18.92
49-72 h	0.06	1.69	66.01	-0.97	2.40	51.99	0.91	0.71	14.02

注:订正效果数据,为逐步回归预报较原模式预报减小的预报误差和提高的准确率,下同。

表5为12时次起报的模式预报及逐步回归订正后预报的  $2\text{ m}$  气温对比。 $72\text{ h}$  内模式预报的温度总体偏低  $0.81^{\circ}\text{C}$  ,绝对误差为  $2.19^{\circ}\text{C}$  ,  $2^{\circ}\text{C}$  内的准确率为  $53.25\%$  。通过逐步回归订正,预报总体偏低的情况有所改善,较模式预报绝对误差减小了  $0.78^{\circ}\text{C}$  ,准确率提高了  $17.24\%$  。从逐  $3\text{ d}$  的预报检验

看,订正效果随预报时效的延迟而降低, $25\sim 48\text{ h}$  内的订正效果最显著,较原模式预报的绝对误差降低了  $0.85$  ,准确率提高了  $18.33\%$  , $1\sim 24\text{ h}$  内的准确率提高了  $17.66\%$  , $49\sim 72\text{ h}$  内订正效果最差,但准确率也提高了  $15.67\%$  。

表5 方案1逐步回归订正前后温度预报检验结果(12时次起报)

时效	逐步回归			原模式			订正效果		
	相对误差 / $^{\circ}\text{C}$	绝对误差 / $^{\circ}\text{C}$	准确率 / $\%$	相对误差 / $^{\circ}\text{C}$	绝对误差 / $^{\circ}\text{C}$	准确率 / $\%$	相对误差 / $^{\circ}\text{C}$	绝对误差 / $^{\circ}\text{C}$	准确率 / $\%$
1-72 h	-0.36	1.41	70.49	-0.81	2.19	53.25	0.45	0.78	17.24
1-24 h	-0.34	1.33	74.59	-1.31	2.09	56.93	0.97	0.76	17.66
25-48 h	-0.58	1.34	70.73	-0.54	2.19	52.40	-0.04	0.85	18.33
49-72 h	-0.16	1.56	65.94	-0.59	2.28	50.27	0.43	0.72	15.67

## 2.2 模式和3套订正方案温度预报误差随预报时效的变化

图1为00时次起报的原模式预报和3套逐步回归订正方案的  $2\text{ m}$  气温的相对误差和绝对误差随预报时效的变化。原模式的气温预报绝对误差随时间的变化具有明显的周期振荡,绝对误差峰值出现在每天的05时和15时前后,即日最低气温和最高气温发生时段,且在15时偏差更大。 $3\text{ d}$  内绝对误差峰值随预报时效延长显著增加。结合相对误差不难发现,原模式气温预报除前0-6 h以外,其余预报时效内预报都是偏低的,其中在夜间(20时至次日08时)偏低明显,偏低最显著的时段依然是05时前后,而其他时段没有显著的偏高或偏低的特征。

分析逐步回归订正后的温度预报结果发现:3套逐步回归订正方案对原模式的气温预报在所有预报时效内都有显著改进,预报偏差随预报时效延长而增加。其中方案1改进效果略优于方案2的,明显优于方案3的(尤其在  $24\sim 72\text{ h}$  预报时效)。方案1

和方案2无论是相对误差还是绝对误差随预报时效的变化趋势几乎完全一致,皆具有明显的周期振荡,绝对误差峰值出现在每天15时前后(最高气温发生时段),有效订正了原模式在05时前后出现的误差峰值,即方案1和方案2的订正效果夜间的(尤其是最低气温)大于白天的(08时至20时)。对比方案1和方案2,方案3的差异主要体现在绝对误差峰值除了在05时前后外,在每天10时前后还存在一个小峰值( $24\sim 72\text{ h}$  预报时效)。

图2为12时次起报的原模式预报和3套逐步回归订正方案的  $2\text{ m}$  气温的相对误差和绝对误差随预报时效的变化。原模式的相对误差在  $72\text{ h}$  内都是负偏差,说明在  $72\text{ h}$  预报时效内气温预报均偏低,且预报偏差随预报时效延长略有增加。绝对误差和00时次起报的一样,具有明显的周期振荡,误差峰值出现在每天的05时和15时前后,且在15时的偏差更大。结合相对误差不难发现,原模式的气温预报在夜间比白天预报偏低更明显。

逐步回归订正后的温度预报显示,3套逐步回归订正方案对原模式的气温预报在所有预报时效内都有显著改进,预报偏差随预报时效延长略有增加。总体来说,方案1优于方案2和方案3,尤其是1-24 h和49-72 h 预报时效内,方案1的改进效果更显著,在13-24 h 预报时效内方案2和方案3改进效果相当,但都次于方案1的,在25-72 h 预报时效内方

案1和方案2改进效果相当,且皆优于方案3的。3套方案的订正效果夜间的大于白天的,其预报偏差具有明显的周期振荡,误差峰值出现在每天的15时前后(最高气温发生时段),误差谷值出现在每天的05时前后(最低气温发生时段)和00时次起报的周期特征是一致的。

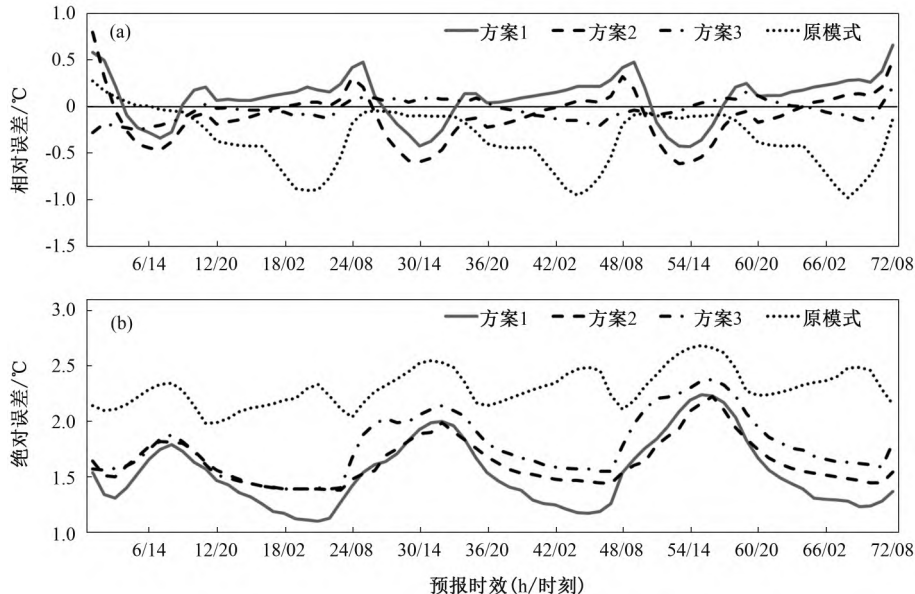


图1 00时次起报的模式预报和3套逐步回归订正方案的2 m 气温的相对误差(a)和绝对误差(b)随预报时效的变化

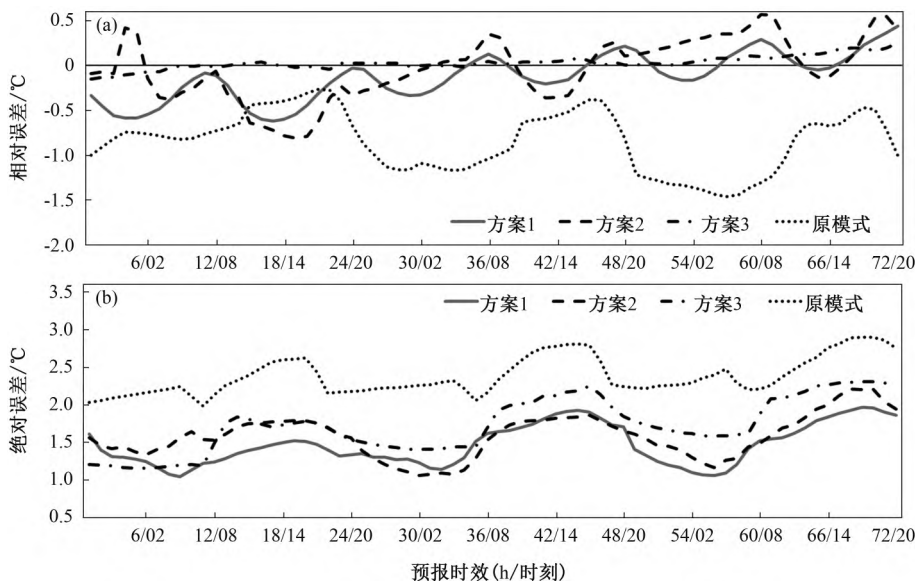


图2 12时次起报的模式预报和3套逐步回归订正方案的2 m 气温的相对误差(a)和绝对误差(b)随预报时效的变化

由于00时次和12时次起报的原模式和3套订正方案的温度预报随预报时效都有着一致的周期特征,因此仅用订正效果最佳的方案1和00时次起报

结果分析订正前后模式的季节性变化。图3为00时次起报的四个季节(3-5月为春季,6-8月为夏季,9-11月为秋季,12月-次年2月,绝对误差为季均值)

模式预报和方案1订正后的2 m 气温的绝对误差随预报时效的变化。总体而言,原模式和方案1的预报结果季节性差异不大,四个季节的预报偏差具有明显的周期振荡。就原模式而言,春、夏、秋三季误差峰值大体出现在每天的05时和15时前后,而冬季误差峰值出现在每天的07时和14时前后,1-48 h 预报时

效内夏季的绝对误差最大,冬季的最小,其他两季预报效果相当。就方案1而言,四季的订正效果没有明显优劣差异,春、夏、秋三季误差峰值大体出现在每天的15时前后,误差谷值出现在每天的05时前后,而冬季误差峰值出现在每天的14时前后,误差谷值出现在每天的07时前后。

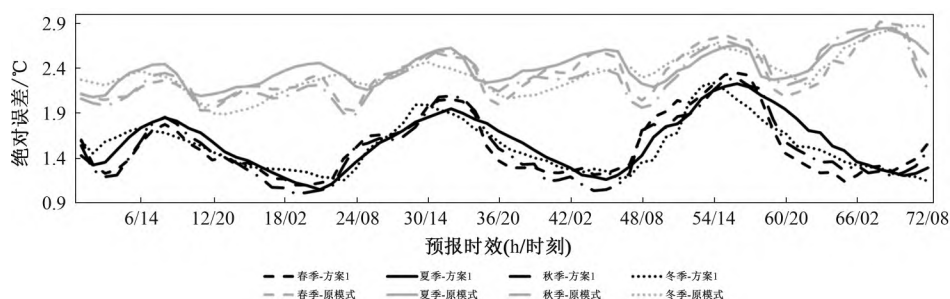


图3 00时次起报的四季模式预报和方案1订正后的2 m 气温的绝对误差随预报时效的变化

### 2.3 3套建模方案的温度预报误差空间分布特征

通过分析原模式与3套建模方案的00时、12时次报的1-72 h 2 m 气温预报相对误差和绝对误差的空间分布,探讨基于逐步回归模型订正前后的华南区域模式在湖南不同地区的预报效果。

图4为原模式预报的2 m 气温1-72 h 平均误差空间分布。原模式00时次起报的2 m 气温在湘东大部(偏高1℃左右)和南岳(偏高6℃以上)偏高以外,其他大部分地区的预报偏低,其中张家界、永州南部和西北部、衡阳西南部、邵阳南部、郴州西部偏低2℃以上(图4a)。原模式12时次起报的2 m 气温平均相对误差空间分布与00时次起报的相似,明显的差异主要体现在湘东,平均相对误差12时次起报的为负偏差,00时次起报的为正偏差(图4b)。原模式00时次起报的2 m 气温平均绝对误差大值区位于南岳,达到6℃以上,张家界、永州南部、衡阳南部和郴州西部平均绝对误差在3℃以上,湘东及湘中一带平均绝对误差相对较小,在2℃以下。总体而言,预报效果东部(平原)的优于西部和南部(山区)的(图4c)。原模式12时次起报的2 m 气温平均绝对误差大值区依然位于南岳,达到5℃以上,其他地区平均绝对误差空间分布和00时次起报结果都非常相似,但误差较00时次起报的更大(图4d)。

图5为00时次起报的3套逐步回归订正方案的2 m 气温1-72 h 时效的平均相对误差和绝对误差空间分布。3套逐步回归订正方案对原模式的气温预报有显著改进,尤其是原模式误差较大的区域。3套逐步回归订正方案的相对误差空间分布差异较小,一致表现为湘西北为正偏差,湘东为负偏差,差

异主要体现在湘东南,方案1在湘东南大部分地区是正偏差,而方案2和方案3是负偏差(图5a-c)。3套方案的平均绝对误差较原模式的明显降低,尤其是南岳地区,改进最为明显(绝对误差减少4℃以上),且其绝对误差空间分布差异非常接近,大体表现为西高东低的分布特征,其中方案1在全省的订正效果最佳,其次是方案2的(图5d-f)。

图6为12时次起报的3套逐步回归订正方案的2 m 气温1-72 h 时效的平均相对误差和绝对误差空间分布。12时次起报的3套逐步回归订正方案同样对原模式的气温预报有显著改进,尤其是南岳。方案1和方案2的相对误差在全省绝大部分地区表现为负偏差,其中方案1的负偏差范围更大,方案3的相对误差空间分布中正偏差范围明显比方案1和方案2的大(图6a-c)。和00时次起报的相比,方案1和方案2中12时次起报的负偏差范围大于00时次起报的,3套方案中正偏差小于00时次起报的。3套方案的平均绝对误差较原模式的明显降低,尤其是南岳地区,改进最为明显,且其绝对误差空间分布差异相似,大体表现为西高东低的分布特征,其中方案1的订正效果最佳,其次是方案2的(图6d-f)。

### 3 结论与讨论

本文针对华南区域模式(新版MARS3 km)2 m 气温预报产品,基于逐步回归模型,依据模式的起报时间、预报时效建立了3套气温订正模型,并分析了模式和3套订正模型气温预报的偏差及不同起报时次和预报时效对模型预报效果的影响,得到以下结论:



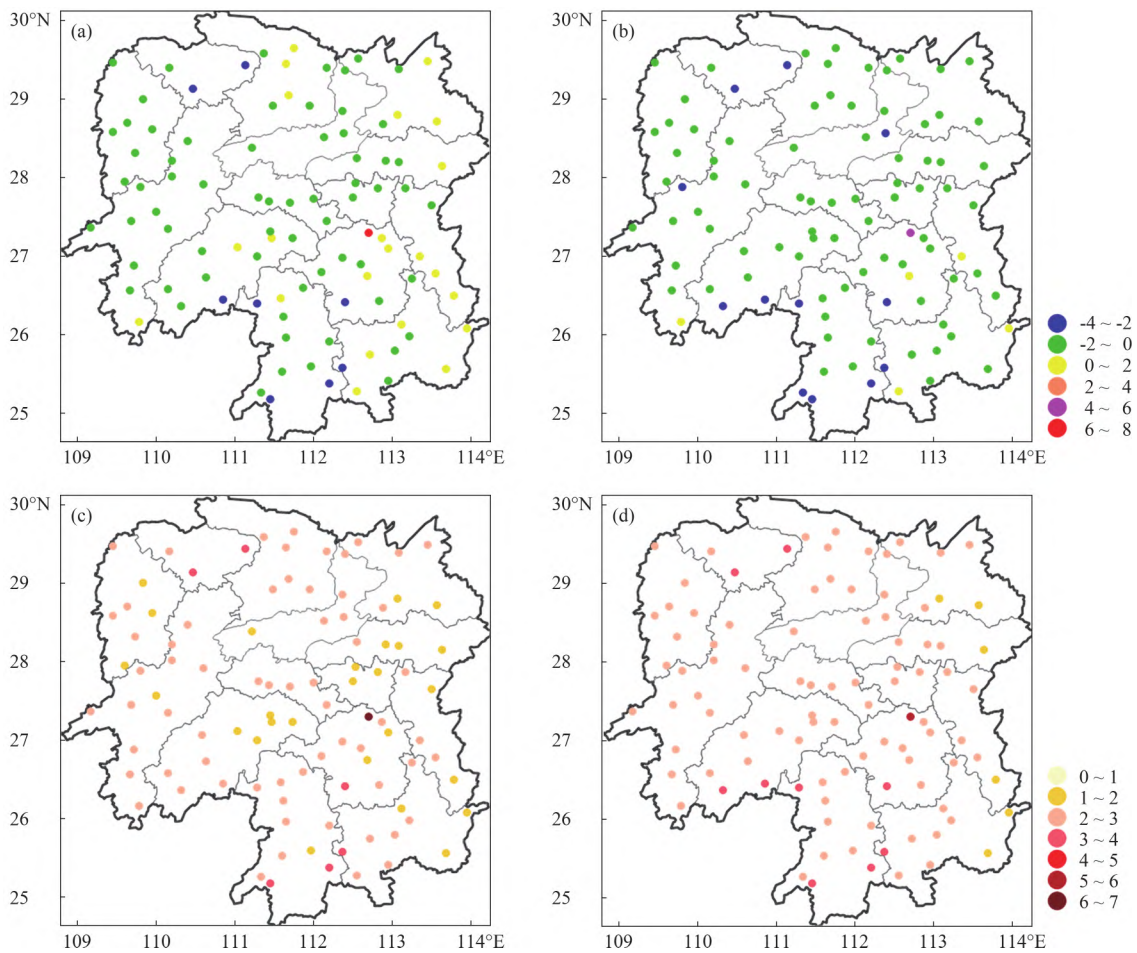


图 4 原模式预报的 2 m 气温 1-72 h 时效的平均误差空间分布

(a) (c) 为 00 时次起报 (b) (d) 为 12 时次起报; (a) (b) 为相对误差 (°C) (c) (d) 为绝对误差 (°C)

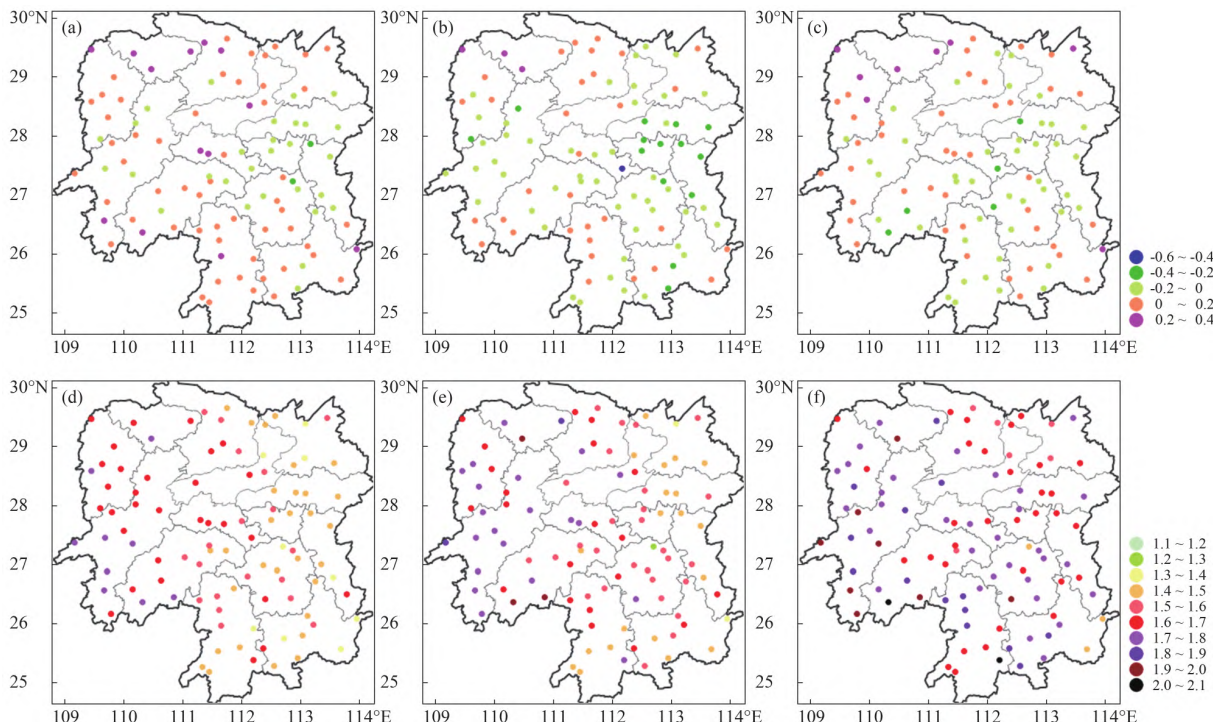


图 5 00 时次起报的 3 套逐步回归订正方案的 2 m 气温 1-72 h 时效的平均误差空间分布

(a) (d) 为方案 1 (b) (e) 为方案 2 (c) (f) 为方案 3; (a) - (c) 为相对误差 (°C) (d) - (f) 为绝对误差 (°C)

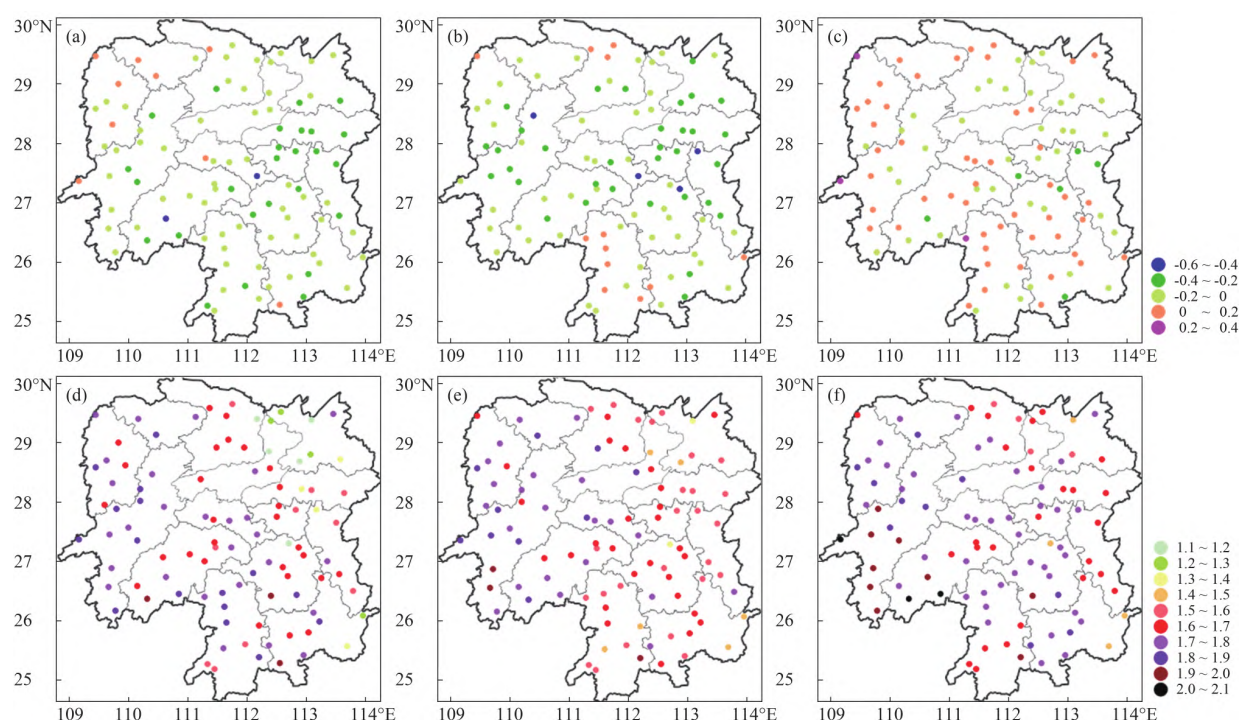


图6 12时次起报的3套逐步回归订正方案的2 m气温1-72 h时效的平均误差空间分布

(a) (d) 为方案1 (b) (e) 为方案2 (c) (f) 为方案3; (a) -(c) 为相对误差(℃) (d) -(f) 为绝对误差(℃)

(1) 华南区域模式尽管分辨率较高,但无法细致反映湖南复杂地形影响导致的区域气温变化特征。00/12时起报的1-72 h时效内的温度预报平均偏差东部的(平原)小于西部和南部的(山区),但总体而言未能有效反映湖南气温的空间变化。

(2) 华南区域模式气温预报在1-72 h预报时效内的准确率为54.08%,预报偏差随时间变化具有明显的周期振荡,偏差大值出现在日最低气温和最高气温发生时段(春、夏、秋三季误差峰值大体出现在05时和15时前后,冬季误差峰值出现在07时和14时前后),且在最高气温时段偏差更大,其中夜间(20时至次日08时)预报明显偏低,谷值在05时前后。1-72 h时效内预报偏差随预报时效延长显著增加。

(3) 3套针对不同起报时次和预报时效的建模方案皆对湖南2 m气温的时空分布特征有较好刻画,有效提升了复杂地形影响区域的气温预报准确率。方案1的气温订正效果总体优于方案2和方案3的,00时起报的1-72 h预报时效内的绝对误差比原模式减小了0.68℃,准确率提高17.14%,12时起报的1-72 h预报时效内的绝对误差比原模式减小了0.78℃,准确率提高17.24%。3套建模方案气温订正准确率随着预报时效的延长而下降。

(4) 3套建模方案2 m气温订正效果在不同起报时次的对比结果表明,1-72 h预报时效内,除了方

案2中49-72 h预报时效及方案3中25-48 h时效外,其余预报时效内3套方案的订正效果都是00时起报的2 m气温预报准确率明显高于12时起报的准确率。

(5) 3套建模方案2 m气温订正效果逐时变化的对比结果表明,方案1和方案2改进效果相当(总体优于方案3),且具有明显的周期振荡,但没有明显的季节差异,最大偏差出现在最高气温发生时段,对夜间气温,尤其是05时前后的气温预报订正效果明显。

逐步回归法尽管极大改善了华南区域模式对2 m气温的预报能力,但是由于逐步回归模型仅仅建立了单站要素的函数关系,并未涉及更大范围的大尺度环流因子,使模拟效果依旧存在偏差。气温变化的影响因子较为复杂,建立湖南地区单站气温与周围大尺度环流因子之间的关系,将是下一步的主要工作。

## 参考文献

- [1] 黄鹤楼, 丁烨毅, 姚日升, 等. 宁波城市化进程对夏季极端气温和高温热浪的影响[J]. 气象与环境科学, 2020, 43(2): 56-62.
- [2] 左璇, 姬兴杰. 均一化前后河南省年平均最高气温和高温日数的对比分析[J]. 气象与环境科学, 2020, 43(4): 36-41.
- [3] 陈昱文, 黄小猛, 李熠, 等. 基于ECMWF产品的站点气温预报集成学习误差订正[J]. 应用气象学报, 2020, 31(4): 494-503.



- [4]沈沉, 赵文灿, 施金海, 等. 多隐层 BP 神经网络在模式预报中的简化应用[J]. 气象与环境科学, 2019, 42(4): 127-132.
- [5]盛春岩, 范苏丹, 荣艳敏, 等. 几种气温客观预报方法对比及最优集成预报研究[J]. 气象, 2020, 46(10): 1351-1361.
- [6]吴乃庚, 曾沁, 刘段灵, 等. 日极端气温的主客观预报能力评估及多模式集成网格释用[J]. 气象, 2017, 43(5): 581-590.
- [7]吕林宜, 王新敏, 栗晗. 华东区域模式对河南“7·19”特大暴雨的数值检验与分析[J]. 气象与环境科学, 2019, 42(1): 101-109.
- [8]张成军, 纪晓玲, 马金仁, 等. 多种数值预报及其释用产品在宁夏天气预报业务中的检验评估[J]. 干旱气象, 2017, 35(1): 148-156.
- [9]盛春岩, 范苏丹, 荣艳敏, 等. 几种气温客观预报方法对比及最优集成预报研究[J]. 气象, 2020, 46(10): 1351-1361.
- [10]李超, 裴薇, 裴小芬, 等. 浙江省温度和相对湿度释用技术及其效果检验分析[J]. 气象科学, 2016, 36(4): 556-561.
- [11]吴启树, 韩美, 郭弘, 等. MOS 温度预报中最优训练期方案[J]. 应用气象学报, 2016, 27(4): 426-434.
- [12]肖瑶, 杜明哲, 徐鹏. 两种订正方法在河南高速公路沿线气温精细化预报中的应用[J]. 干旱气象, 2020, 38(6): 1023-1030.
- [13]赵声蓉. 多模式温度集成预报[J]. 应用气象学报, 2006, 17(1): 52-58.
- [14]朱文达, 张媛, 杨静, 等. 2015 年华南区域模式贵州区域 2 m 气温预报偏差分析[J]. 中低纬山地气象, 2018, 42(2): 28-35.
- [15]马旭林, 时洋, 和杰, 等. 基于卡尔曼滤波递减排平均算法的集合预报综合偏差订正[J]. 气象学报, 2015, 73(5): 952-964.
- [16]白永清, 林春泽, 陈正洪, 等. 基于 LAPS 分析的 WRF 模式逐时气温精细化预报释用[J]. 气象, 2013, 39(4): 460-465.
- [17]门晓磊, 焦瑞莉, 王鼎, 等. 基于机器学习的华北气温多模式集合预报的订正方法[J]. 气候与环境研究, 2019, 24(1): 116-124.
- [18]陈迪, 陈豫英, 马金仁, 等. 不同时间尺度的 MOS 方法对宁夏气温预报的影响[J]. 干旱区地理, 2019, 42(1): 94-102.
- [19]罗聪, 曾沁, 高亭亭, 等. 精细化逐时滚动温度预报方法及检验[J]. 热带气象学报, 2012, 23(4): 395-401.
- [20]齐铎, 刘松涛, 张天华, 等. 基于格点的中国东北中北部 2 m 温度数值预报检验及偏差订正[J]. 干旱气象, 2020, 38(1): 81-88.
- [21]王婧, 徐枝芳, 范广洲, 等. GRAPES-RAFS 系统 2 m 温度偏差订正方法研究[J]. 气象, 2015, 41(6): 716-726.
- [22]王秀娟, 陈长胜, 冯旭, 等. 一阶卡尔曼滤波方法对 EC 集合预报气温的订正[J]. 气象灾害防御, 2019, 26(1): 34-38.
- [23]李刚, 彭芳, 孔德璇, 等. 华南 3 km 高分辨率区域模式气温预报在贵州的质量检验[J]. 中低纬山地气象, 2019, 43(6): 1-8.
- [24]康岚, 冯汉中, 屠妮妮, 等. Grapes 模式预报西南地区夏季 2 m 温度的检验评估[J]. 高原山地气象研究, 2009, 29(2): 26-32.
- [25]李佰平, 智协飞. ECMWF 模式地面气温预报的四种误差订正方法的比较研究[J]. 气象, 2012, 38(8): 897-902.
- [26]王敏, 李晓莉, 范广洲, 等. 区域集合预报系统 2 m 温度预报的校准技术[J]. 应用气象学报, 2012, 23(4): 395-401.
- [27]李超, 李明华, 周凯, 等. 基于 KNN 回归算法的浙江省温度预报改进研究[J]. 气象与环境科学, 2022, 45(1): 81-89.
- [28]殷悦, 吴利红, 严睿恺, 等. 区域气象站日气温极值延长方法研究与误差分析[J]. 气象与环境科学, 2023, 46(4): 95-103.
- [29]赵宁坤, 张秀年, 孙俊奎, 等. 高分辨率区域模式降水预报在云南的检验[J]. 暴雨灾害, 2021, 40(1): 78-86.
- [30]李刚, 彭芳, 孔德璇, 等. 华南 3 km 高分辨率区域模式气温预报在贵州的质量检验[J]. 中低纬山地气象, 2019, 43(6): 1-8.
- [31]朱文达, 陈子通, 张艳霞, 等. 高分辨地形对华南区域 GRAPES 模式地面要素预报影响的研究[J]. 热带气象学报, 2019, 35(6): 801-811.
- [32]王丹, 高红燕, 马磊, 等. SCMO 温度精细化指导预报在陕西区域的质量检验[J]. 气象科技, 2014, 42(5): 839-846.

## Research on Interpretation and Application of Temperature Forecast Products of Regional Model in South China Based on Stepwise Regression Model

Lan Mingcai<sup>1,2</sup>, Tang Jie<sup>1,2</sup>, Zhou Li<sup>1,2</sup>, Chen Jingjing<sup>1,2</sup>, Tan Ping<sup>3</sup>, Liu Qingzhu<sup>4</sup>

(1. Hunan Meteorological Observatory, Changsha 410118, China;

2. Key Laboratory of Preventing and Reducing Meteorological Disaster of Hunan Province, Changsha 410118, China;

3. Hunan Meteorological Service Center, Changsha 410118, China;

4. Jimei District Meteorological Office, Xiamen 361021, China)

**Abstract:** Based on the 2 m temperature prediction products of South China mesoscale weather prediction model from 2017 to 2019, three sets of model construction schemes were designed using 25 significant factors to establish the stepwise regression correction model of 97 national basic stations. The deviation between the revised prediction of the three sets, the models and the actual situation were analyzed. The results show that: (1) although the regional model of South China has high resolution, it could not

reflect the characteristics of regional temperature change caused by the complex terrain of Hunan. The deviation in mountainous areas is greater than that in plain areas. The accuracy is 54.08% within 1–72 hours. The prediction deviation increases significantly with the extension of prediction time, which has obvious periodic oscillation. The maximum deviation occurs with daily maximum/minimum temperature. (2) Based on the stepwise regression model, three sets of temperature correction models are established according to the starting time and prediction time of the model. Scheme 1 is the modeling not distinguishing starting time and prediction time, scheme 2 is the modeling distinguishing starting time and not distinguishing forecasting time, and scheme 3 is the modeling distinguishing both starting time and prediction time. The three sets of modeling schemes all have a good description of the temporal and spatial distribution characteristics of 2 m air temperature in Hunan. It has certain correction ability for the temperature forecast in the area affected by complex terrain. However, with the extension of prediction time, the correction effect decreases. The overall correction effect at 00:00 start time is better than that at 12:00. (3) The 2 m temperature forecast of the three modeling schemes has the characteristics of periodic oscillation, and no obvious seasonal difference. The maximum deviation occurs with maximum temperature. The correction effect is obvious for the night temperature, especially before and after 05:00. The temperature correction effect of scheme 1 is better than that of scheme 2 and scheme 3. The absolute error within 1–72 h prediction time is reduced by 0.68 °C, and the accuracy is improved by more than 17%.

**Key words:** stepwise regression; South China regional model; 2 m temperature; temperature correction