文章编号:1674-7070(2014)05-0449-10

黄惠镕[1]　郭品文[1]

基于统计降尺度方法的夏季降水精细化预报

摘要

利用1999—2009年安徽省淮河以南地区60个县市站夏季逐日降水资料和安庆市探空站逐日资料,研究了中低层不同风向配置下局地降水与大尺度降水场之间的关系,以3种不同预报对象及相应的预报因子分别采用神经网络和线性回归方法设计6种预报模型对观测资料进行逼近和优化,从而实现空间降尺度.分析对比6种预报模型46站逐日降水量的拟合和预报效果,结果表明:采取相同的预报对象及预报因子的BP神经网络模型在拟合和预报效果上均好于线性回归模型,可见夏季降水场之间以非线性相关为主;神经网络模型预报结果同常用的Cressman插值预报相比,能很好地反映出降水的基本分布及局地特征;预报对象为单站降水序列的神经网络模型在以平原、河流为主要地形的区域预报效果较好,预报对象为REOF主成分的神经网络模型则在山地和丘陵地形区域预报效果较好.

关键词

统计降尺度;日降水;BP神经网络;旋转主成分分析(REOF)

中图分类号 P456.8

文献标志码 A

收稿日期 2012-03-19

资助项目 公益性行业(气象)科研专项(GYHY201006017)

作者简介

黄惠镕,女,硕士生,主要从事天气预报理论与方法研究.ytyy31@sina.com

郭品文(通信作者),男,教授,研究方向为短期气候预测.guo@nuist.edu.cn

1　南京信息工程大学　气象灾害省部共建教育部重点实验室,南京,210044

0　引言

由于目前大尺度数值模式输出的空间分辨率较低,缺少区域信息,很难对区域气象要素做精确的预测.有2种方法可以弥补这一不足,一是发展更高分辨率的全球气候模式(GCM)或天气预报模式,二是利用模式预报技巧较高的大尺度信息给出降水、温度等小尺度预报即借助降尺度方法,而GCM及天气预报模式计算量较大,对分辨率提高潜力有限,因此降尺度方法更为有效[1].目前国内外主要有2种降尺度方法:动力降尺度法和统计降尺度法.使用动力降尺度方法,可以内嵌入全球模式,或者利用全球模式的输出作为边界条件单独运行.动力降尺度具有物理意义明确、不受观测资料影响和面向覆盖区域的所有格点等很多优点,但它计算量大,模拟和配置不便.统计降尺度方法的基本原理在于采用统计经验的方法建立大尺度气象变量与区域气象变量之间线性或非线性联系[2],其优点在于它和现有模式产品相配合,能较好地解决现有动力降尺度法的不足,是提高气象要素的精细化预报水平相对直接有效的手段,在季节尺度的气候预测上也得到了广泛应用[3-8],而在短期气象要素预报领域的研究和应用较少.

常用于降尺度的气象要素有降水、气温、气压、湿度、风速、蒸散发量等,其中降水在时间与空间上高度不连续,是最难模拟的变量,同时它也是对洪涝模拟最为敏感的驱动参数,因而成为降尺度研究的主要气象要素,同时也是天气预报业务中重要的预报对象.我国气象要素预报目前为精细到县站级别的站点预报,在实践中雨量观测站点的布设受到地形、人力、财力等因素的限制,直接观测数据无法精确表达降水量的时空分布,通常需要采用空间插值方法对站点观测数据进行插值,以获取整个区域内降水量的总体分布数据,从而实现降水资料的空间降尺度.前人研究[9]表明不同的插值方法对区域气象要素数据计算的误差也有所差异,数据密度、数据分布和空间异质性是影响插值精度的主要因素.当数据密度足够大时,各种插值方法用于同一组气象数据的插值结果差异不大,但在数据分布稀疏和差异显著时,插值方法的选用及其计算参数的设定就尤为重要,因此采用简单插值的统计降尺度方法局限很大.

各种线性和非线性函数一般只能生成单站点的降尺度,基于降水的精细化预报业务需求,研究新的空间降尺度技术有重要意义.本

文利用安徽省站点资料研究逐日降水量的降尺度技术同常用的 Cressman 客观分析方法的插值作对比分析,为短期气象要素的多站点统计降尺度预报拓宽新的思路和方法,具有较强的实际业务应用价值.

1 资料选取及方法

1.1 资料选取

安徽省每年的降水大多集中在夏季,而夏季日降水量复杂多变.淮河以北地区主要以平原地形为主,淮河以南地区地形则复杂多变,致使南部地区降水更难以预测,因此淮河以南降水的精细化预报显得尤为重要.本文选取的是安徽省淮河以南地区 60 个县市站 1999—2009 年夏季(6—8 月)的逐日降水量资料以及同时段的安庆站逐日探空资料.

1.2 EOF 方法及旋转 EOF 方法

采用 EOF 展开的主成分不仅浓缩了初选因子的预报信息,同时各主成分之间是相互正交的,消除了因子之间的复共线性,减少了重复信息和噪音.对一个预报因子群 X 进行经验正交函数(Empirical Orthogonal Function,EOF)分解,可分解成时间函数 Z 和空间函数(特征向量)V 两部分,其数学表达式为

$$X = VZ. \tag{1}$$

由于空间函数(特征向量)随时间变化量很小,所以初选预报因子随时间变化的信息都集中反映在时间函数(主成分)的变化上.每一个新的综合预报因子变量是初选因子变量的线性组合,各个综合因子变量的特征值明确表示了它对初选因子的贡献大小.

旋转主成分分析(Rotated Empirical Orthogonal Function,REOF)是在传统主成分分析(EOF)的基础上再做旋转.本文采用 Horel 使用的方差最大正交旋转法,也是气候分析和诊断经常使用的方差最大正交旋转法,其基本原理详见文献[10].它可以将空间场上与主成分相关的高值区集中在一个较小的范围,从而更容易识别空间型,旋转后每一个空间点只与一个主成分存在高相关,旋转后的特征场比旋转前在时间上更稳定.因此气象上经常采用 REOF 的方法对气象要素场进行分析,它使原来的特征向量结构简化.REOF 分析不仅可以很好地反映不同地域的变化,而且可以反映不同地域的相关分布状况.

1.3 人工神经网络方法

人工神经网络是一个可自动提取一组预报变量和另一组自变量之间非线性关系的数据处理系统,该网络的建立过程称为人工神经网络的训练过程,是用自适应算法递归迭代求解因变量与自变量之间的非线性关系,训练后的神经网络模型可用来估算或预报预测变量.它具有许多优良性能,如自组织性和自适应性能力、记忆联想能力、容错性等[11-12].神经网络的模型很多,本文使用的 BP 神经网络(Back Propagation Neural Network)模型是目前应用最广泛,且具有成效的、算法较成熟的一种.该模型实际计算输出为

$$\hat{y}_j = f\left(\sum_{i=1}^{p} b_i w_{ij} + \gamma_j\right), \tag{2}$$

式中 b_i 是输入层到隐含层新的激活值,w_{ij} 是连接权系数,其初始时刻为一组给定的随机小量,γ_j 是输出层单元阈值,f 取 Sigmoid 函数.有关 3 层前馈网络模型[11] 的算法为

$$f(x) = 1/(1 + e^{-x}). \tag{3}$$

1.4 Cressman 插值法

Cressman 客观分析插值法插值公式如下:

$$V_S = \sum_{i=1}^{N} V_i w_i \Big/ \sum_{i=1}^{N} w_i, \tag{4}$$

w_i 为第 i 个站点的要素值和该站点对站点 S 影响的权重函数:

$$\begin{cases} w_i = (R_0^2 - R_i^2)/(R_0^2 + R_i^2), & R_i \leqslant R_0, \\ w_i = 0, & R_i > R_0, \end{cases}$$

N 为可能对站点 S 造成影响的站点总数,V_S 为被插值点值,R_i 和 R_0 分别为第 i 个站点到插值站点 S 的水平距离和站点 S 在插值时考虑的影响半径.

2 预报量和预报因子分析

2.1 预报量及预报因子选取

在淮河以南地区 60 个县市站选择具有区域代表性及分布较均匀的 14 站(寿县、蚌埠、滁县、六安、桐城、合肥、巢湖、芜湖、芜湖县、安庆、宁国、屯溪等)作为大尺度降水场的代表,其余 46 站代表局地降水场尺度(站点分布如图 1 所示).

本文主要讨论大尺度场降水场和局地降水之间的关系.在不同天气系统下,夏季降水量的大小及分布都有所不同,进一步细化到局地尺度上,由于地形的直接和间接影响,局地降水量分布上也有很大的差异.700 hPa 风向即受地形影响又受天气系统的影响,首先利用 700 hPa 风向对夏季降水资料进行细分,统计了 1999—2009 年夏季安庆探空站 700 hPa 层次下各风向(静风<2 m/s)频率.11 a(33 个月)共 1 012 d 的时间里,出现天数较多的有 3 个风向,其中

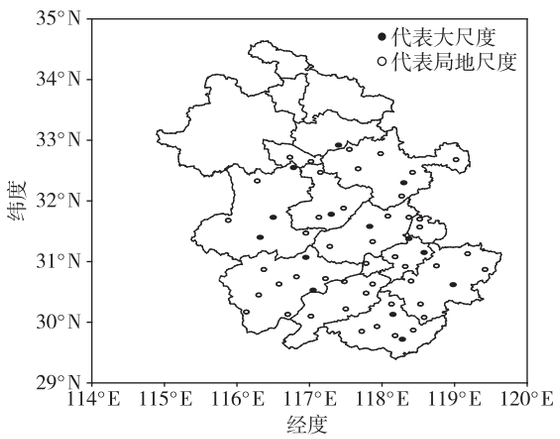


图 1 淮河以南地区选取站点分布

Fig. 1 Distribution of 60 selected stations.

南风有 131 d,西南风有 315 d,西风有 226 d.根据大尺度降水场分为有雨日和无雨日,其中有雨日定义为 14 站中只要有一站降水不为 0 的情况.在不同风向下,将 1999—2008 年及 2009 年 2 个样本分别挑出有雨日,利用 1999—2008 年的有雨日作为训练样本进行模型的构建,2009 年的有雨日资料作为检验样本.

由表 1 可以看到北风及东北风出现频数较少,降水日数少,因此不考虑这 2 个风向.考虑到西南风及西风样本数较多,进一步利用中层风向资料进行划分,500 hPa 风向以西风和西南风为主,下文将 700 hPa 为西南风,中层风向为西南风和西风分别称为西南风 1、西南风 2,西风同样划分为西风 1 及西风 2.这 4 个风向下的训练样本长度为 105、102、29、137,2009 年有雨日的天数分别为 11、9、5、9 d.

表 1 安庆探空站夏季风向分布及对应风向下的大尺度有雨日

Table 1 The summer wind days and corresponding large-scale rainy days recorded by Anqing sounding station d

风向	1999—2008 年		2009 年	
	总日数	有雨日	总日数	有雨日
静风	37	29	5	5
北风	6	2	1	1
东北风	37	20	6	5
东风	72	47	3	2
东南风	99	74	6	3
南风	120	93	11	11
西南风	286	248	29	26
西风	201	187	25	23
西北风	62	44	6	6

2.2 降尺度方案设计

在 9 种不同风向配置下,首先对大尺度降水场的 11 a 降水资料做 EOF 分析,得到其主要的降水空间模态和时间系数,有效地对原来的数据集进行压缩和降维.预报量的选取有 3 种方法:1)普通的 PP 预报都是从单站出发,利用单站降水时间序列作为预报量;2)对 46 站降水做 EOF 分析,选取前几个累计方差贡献率≥90%的特征向量的时间系数作为预报变量进行模型构建;3)考虑到小尺度的降水场局域性较强,将 REOF 方法应用到局地降水场,选取 EOF 累计方差贡献率达到 90%的前几个主成分进行旋转,得到新的预报变量.针对不同预报变量选取相关系数通过 0.05 显著性检验的大尺度降水场主成分作为预报因子,利用 BP 神经网络和多元线性回归方法构建模型来探讨预报量与预报因子之间的线性或非线性关系,最后用拟合和预报的时间系数及对应的空间结构场进行 EOF、REOF 重构,还原后可得到降水量.将 6 种模型分别称为神经网络模型 1、神经网络模型 2、神经网络模型 3、线性回归模型 1、线性回归模型 2、线性回归模型 3,文中分别简称为 BP1 模型、BP2 模型、BP3 模型、LR1 模型、LR2 模型、LR3 模型.

在 9 个风向中西南风整体配置下的日数最多,下面以西南风 1 为例讨论其配置下建模过程中预报因子的选取.西南风 1 下训练样本长度为 105,对其做 EOF 展开,得到前 11 个时间序列作为 BP2 模型、LR2 模型的预报分量,同理用相同训练样本做旋转 EOF 得到前 11 个主成分作为 BP3 模型、LR3 模型的预报分量.西南风 1 下 46 站 EOF 前 3 个主成分的方差贡献率分别为 51.12%、15.62%、5.74%,表 2 给出了主成分对应预报因子的相关系数及方差贡献率.从表 2 可以看出:当预报变量为 46 站 EOF 主成分时,其与对应的大尺度 EOF 主成分相关较为紧密,相关系数分别达到了 0.94、0.74、-0.45,这体现出利用风向配置进行降水资料细化后,大尺度降水场和局地降水相关性强.但前 3 个主成分对应的预报因子方差贡献率总和分别为 66.41%、28.08%、75.13%,预报因子个数为 2、4、7,而 REOF 前 3 个主成分所对应的通过显著性检验的预报因子(表略)个数分别为 6、5、6,所占的方差贡献率之和较大分别为 89.31%、28.66%、69.08%.根据预报变量的方差贡献率与预报因子方差贡献率总和对比来看,REOF 的主成分选取了大尺度场相对更多的浓缩综合因子.EOF 和 REOF 的第 4 至 11 个主成分挑

选预报因子的方法相同.

表 2　西南风 1 下 46 站 EOF 主成分选取的预报因子的相关系数和方差贡献率

Table 2　The correlation coefficients and the variance contributions of predictors selected by EOF principal components in southwest wind 1 of 46 stations

预报对象(46 站)	所选预报因子(14 站)	方差贡献率/%	方差贡献率总和/%	相关系数
EOF 第 1 主成分	EOF 第 1 主成分	53.13	66.41	0.94
	EOF 第 2 主成分	13.28		0.17
EOF 第 2 主成分	EOF 第 2 主成分	13.28	28.08	0.74
	EOF 第 3 主成分	7.52		0.41
	EOF 第 4 主成分	6.22		0.31
	EOF 第 10 主成分	1.06		0.17
EOF 第 3 主成分	EOF 第 1 主成分	53.13	75.13	0.22
	EOF 第 3 主成分	7.52		−0.45
	EOF 第 4 主成分	6.22		0.38
	EOF 第 5 主成分	5.11		0.24
	EOF 第 6 主成分	4.05		0.40
	EOF 第 7 主成分	3.19		0.31
	EOF 第 8 主成分	1.69		0.30

3　预报模型和检验分析

3.1　模型拟合效果分析

本文所采用的神经网络是单隐层网络结构,输入层和隐层神经元采用 f 取 Sigmoid 函数,输出层则采用线性函数作为传递函数.对于隐层节点数的选取,按照式(5)进行计算:

$$n_1 = \sqrt{m + n} + a. \tag{5}$$

每个预报变量由于选择的因子数不同,所以隐节点数也不同,其中 n_1 为隐层单元数,m 为输出层单元数,n 为输入层单元数,a 为 1~10 之间的常数,本试验中 a 的选择从 1~10 进行循环.神经网络参数预报模型的输入节点数取预报因子个数,输出节点数取 l,学习矩阵长度根据风向配置日数(即样本长度)的变化而变化,如西南风 1 第 1 主成分网络结构为 2-n_1-1,其中 n_1 参照式(5)选取变化,学习矩阵长度为 105.学习因子和动量因子分别取 0.9 和 0.7,训练函数取误差梯度下降函数.当网络进行学习训练 5 000 次时,误差函数趋于稳定,保留训练结束时平均绝对误差相对最小时对应的 a 值.此时网络模型确定的连接权系数阀值可以得到相应的预报模型.为了进行客观定量分析,采用以下 2 种指标进行分析.

1) 平均绝对误差:$E_{\mathrm{MA}} = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i |$.

2) Pearson 相关系数:

$$R_{\mathrm{P}} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^2}}.$$

表 3 给出了不同风向下及整体夏季 6 种预报模型对日降水量历史样本的拟合效果.结果表明,在 700 hPa 7 个风向下,神经网络的 3 种预报模型拟合在 MAE、PR 指标上均好于线性回归模型,其中又以 BP1 模型和 BP3 模型的拟合效果较好.在 8 个风向下,静风降水拟合平均绝对误差最小,BP1 及 BP3 模型达到最小,分别为 1.95、2.23 mm,PR 系数分别为 0.78 和 0.73,西南风下降水拟合 MAE 最大,BP1 和 BP3 模型分别为 4.83、4.95 mm,PR 系数均为 0.87,同其他 4 种模型相比,这 2 种模型的拟合效果是较好的.BP1 模型和 BP3 模型的夏季拟合平均绝对误差分别为 3.79、3.99 mm,BP2 模型为 4.77 mm,而线性回归模型整体夏季误差都在 5 mm 以上,PR 分别为 0.90、0.85、0.89、0.80、0.79、0.79.可见从 2 种指标上来看,神经网络模型的拟合精度比线性回归模型更高,其中以 BP1 模型、BP3 模型拟合效果较好.

表 3　6 种降尺度预报模型拟合效果评价指标比较

Table 3　Comparison of fitting evaluation results by 6 kinds of downscaling prediction models

风向	区域均值	BP1 模型	BP2 模型	BP3 模型	LR1 模型	LR2 模型	LR3 模型
静风	MAE	1.95	3.11	2.23	3.13	3.40	3.15
	PR	0.78	0.66	0.73	0.72	0.62	0.63
东风	MAE	1.87	2.72	2.23	3.01	3.13	3.17
	PR	0.82	0.75	0.79	0.72	0.70	0.70
东南风	MAE	2.11	3.18	2.26	3.97	3.81	4.08
	PR	0.87	0.76	0.87	0.65	0.62	0.63
南风	MAE	2.66	3.66	3.03	4.18	4.26	4.58
	PR	0.87	0.80	0.82	0.72	0.71	0.71
西南风	MAE	4.83	5.68	4.95	6.33	6.08	6.32
	PR	0.87	0.83	0.87	0.77	0.77	0.76
西风	MAE	2.66	3.66	3.03	4.18	4.26	4.58
	PR	0.87	0.80	0.82	0.72	0.71	0.71
西北风	MAE	3.65	4.32	4.02	4.82	4.67	5.13
	PR	0.90	0.85	0.87	0.85	0.82	0.84
夏季	MAE	3.79	4.77	3.99	5.34	5.31	5.58
	PR	0.90	0.85	0.89	0.80	0.79	0.79

3.2 独立样本预报检验

3.2.1 西南风总降水量预报检验

评价一个预测模型的优劣分析其拟合效果是一个方面,更重要的是检验模型预报能力.本文利用2009年资料对6种预报模型进行预报检验.

图2给出了西南风23 d总降水量观测值和预报值的空间分布.由图2a可以看出,降水主要呈南北分布,沿淮地区降水偏少,降水总和普遍少于150 mm,但在凤阳有一小范围的降水大值区.南部地区降水偏多,并在长江中下游出现大值区,庐江台站出现降水最大值395.8 mm,降水等值线由中心位置向北向南减少,江淮东部降水普遍偏多在250~400 mm之间.西南风的地形作用,则体现在皖南山区、大别山区的迎风坡上的降水多于背风坡的.BP1模型和BP3模型较好地模拟出降水的大致形势和大值中心,BP1模型对于江淮东部区域的降水总和预报效果较好,但是在沿江西部区域和皖南山区迎风坡上存在降水预报偏小的现象,庐江台站大值中心预报值为417.4 mm,预报偏差为21.6 mm;BP3模型没有预报出凤阳的降水大值,对于庐江的大值降水预报值为358.4 mm,预报偏差为37.4 mm,预报误差相对较大,其对皖南山区的降水预报形势较为准确.BP2模型和LR1模型对降水的预报值能大致反映出降水中间大值区型,针对在250 mm以上区域降水预报效果不好,大值中心和分布走向均不能很好地预报出来.LR2模型在庐江附近及沿江西部和皖南山区均存在预报总和值偏小的情况,LR3模型的预报效果最为不好,降水预测值集中在150~250 mm之间,插值预报同BP1模型一样能反映出降水中间型,局地降水特征没有得到较好体现,其预报的庐江23 d降水量总和为250.0 mm,与真实值相比偏差达到145.8 mm,在整个长江流域的降水大值区域出现了较大的偏差.

3.2.2 夏季降水预报检验

3.2.2.1 夏季总降水量预报

前面只是分析了西南风23 d的预报情况,其余6个风向下总共有46 d也需进行模型的预报检验.下面对2009年夏季69 d有雨日预报情况进行分析.

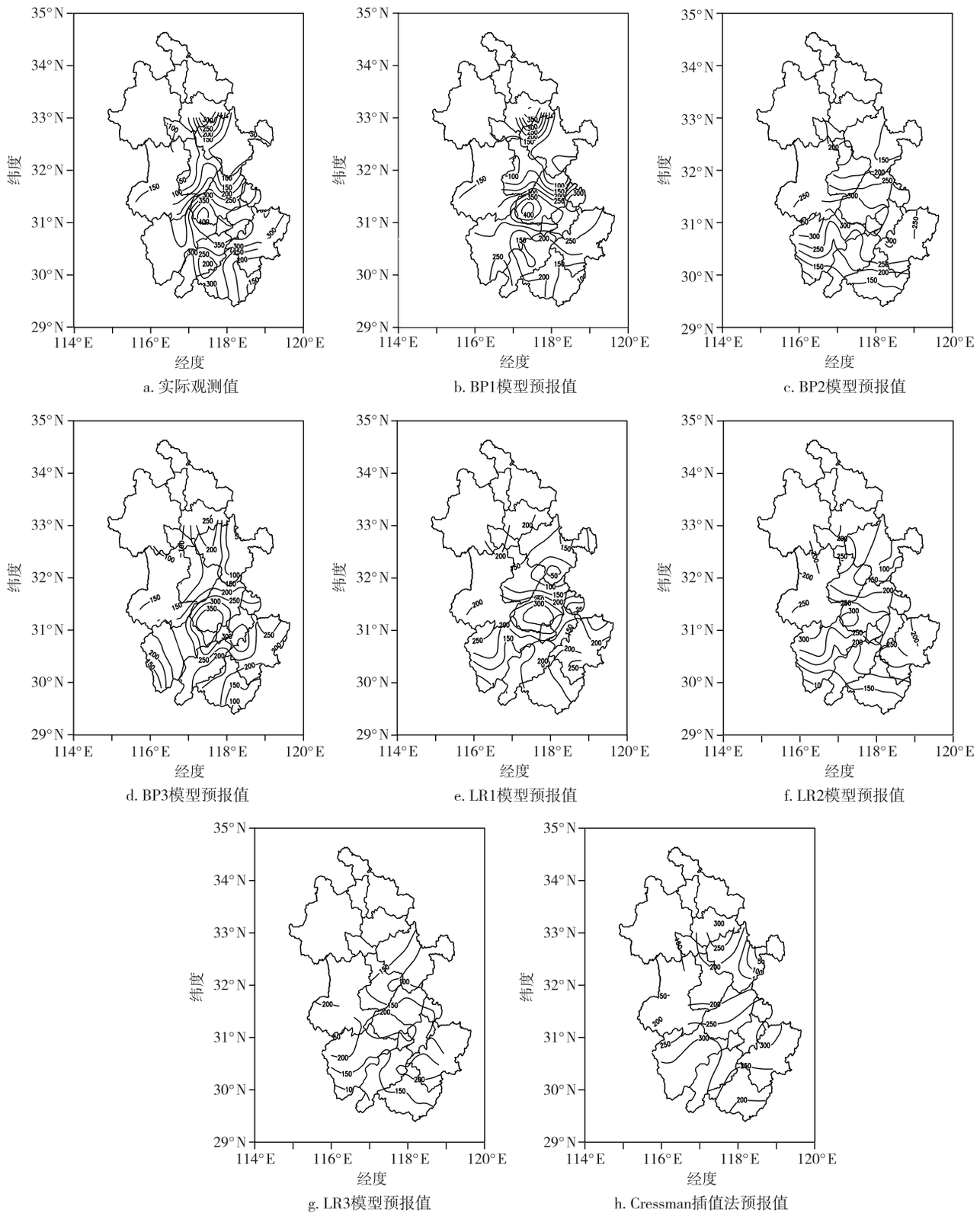
从图3a中可以看到沿淮区域夏季观测降水总和在350~500 mm之间,在凤阳地区有一小的降水大值中心,此处以丘陵地形为主,同沿淮区的平原地形相比降水较多;江淮区域降水普遍较多(超过600 mm),在其中部地区有一大于750 mm的降水极大值区,皖南山区在黄山市附近也有一个降水大值区,宣广郎一带降水偏少(小于500 mm).前3种神经网络模型都反映了基本降水趋势,但BP2模型则未将真实的大值中心预报出来,反而在石台县附近预报出虚假的大值中心,而BP1模型和BP3模型对降水极大值中心的预报效果很好,BP3模型预报的实际大于600 mm的区域较为准确,但比实际中少预报出沿江西部和皖南山区的降水大值,BP1模型在江淮区域的其余地区预报值偏小,但对于皖南山区降水预报效果较好.其余3个线性回归模型预报的江淮区域降水普遍减少200 mm,出现较大的预报偏差,而且3种模型在石台县附近都预报出了降水大值中心,这与实际降水极值中心不符合.插值预报没有很好地反映具体的局地降水细节,整体降水等值线较为平滑,预报大值中心区域往南偏移.

3.2.2.2 夏季降水预报误差指标评价

分析3种神经网络预报模型与插值方法的MAE空间分布(图4)可以看出,BP1模型在合肥以北降水较少区域误差在3~4 mm之间,而在降水较大的区域,MAE在4~6 mm之间,在大值中心区域有6 mm以上的平均绝对误差,BP3模型则表现在4~6 mm区域更多,两者在青阳县附近有大于6 mm以上的小范围预报误差大值区.BP2模型大部分区域的MAE均在6 mm以上,且在凤阳、大别山、及沿江东部区域都有大于7 mm的预报误差.其余3种线性模型下(图略),MAE在6 mm以上区域较大,并在局地出现了9 mm的大值误差中心区,插值预报(图4d)在全区域以4~6 mm为主,局部有超过10 mm以上的大值误差中心.

进一步分析这BP1和BP3模型及插值预报的MAE标准差分别为1.13、1.10、1.43 mm,最大值分别为7.36、7.38、10.1 mm,最小值分别为2.43、1.74、2.52 mm,降水极大值(无为站771.1 mm)对应的分别是4.57、5.69、5.57 mm,降水极小值(肥西站295.1 mm)对应的分别是2.43、3.46、2.52 mm.可见BP1模型对于夏季降水极值的预报情况预报效果较好,但是就MAE的最大和最小值及标准差来看,BP3模型预报在全区范围内更加稳定可靠,而两者在降水总和量预报上及误差空间分布上均优于普通常用的Cressman插值法.

将46站划分为沿淮区、沿江西部区、大别山区、江淮东部区、皖南山区5块区域,分区域探讨Cressman插值及6种预报模型预报误差.通过表4可



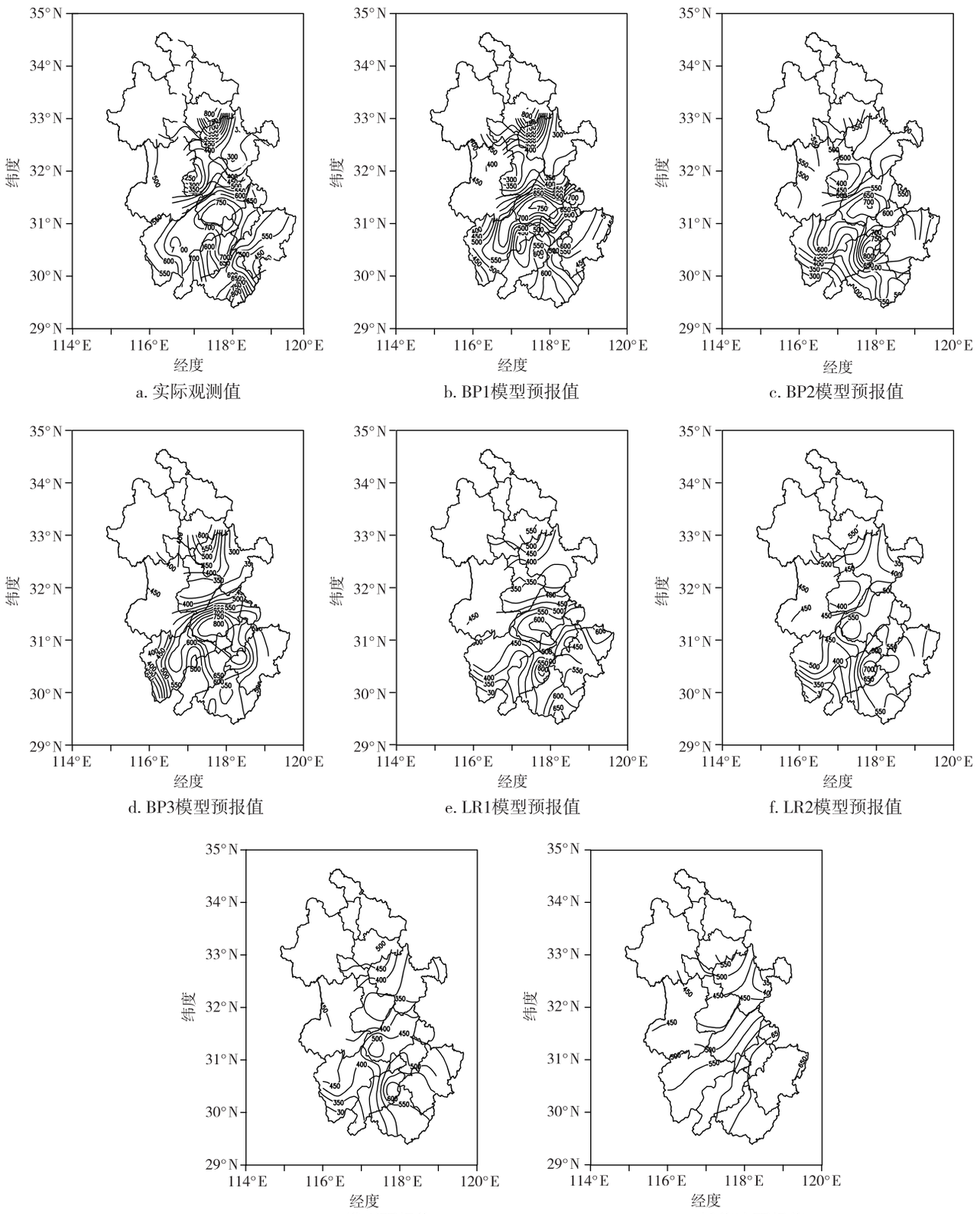
a. 实际观测值　b. BP1模型预报值　c. BP2模型预报值

d. BP3模型预报值　e. LR1模型预报值　f. LR2模型预报值

g. LR3模型预报值　h. Cressman插值法预报值

图 2　西南风 23 d 总降水量实测值和降尺度模型预报值的比较

Fig. 2　Comparison between the observed and predicted precipitation in 23 days of southwest wind by downscaled models with (a) for observation data,and (b) to (h) for forecasting results by BP1,BP2,BP3,LR1,LR2,LR3,and Cressman interpolation method,respectively



a. 实际观测值　b. BP1模型预报值　c. BP2模型预报值

d. BP3模型预报值　e. LR1模型预报值　f. LR2模型预报值

g. LR3模型预报值　h. Cressman插值法预报值

图 3　夏季总降水量实测值与降尺度模型预报值的比较

Fig. 3　Comparison betweer the observed and predicted precipitation for total precipition in summer by downscaled models

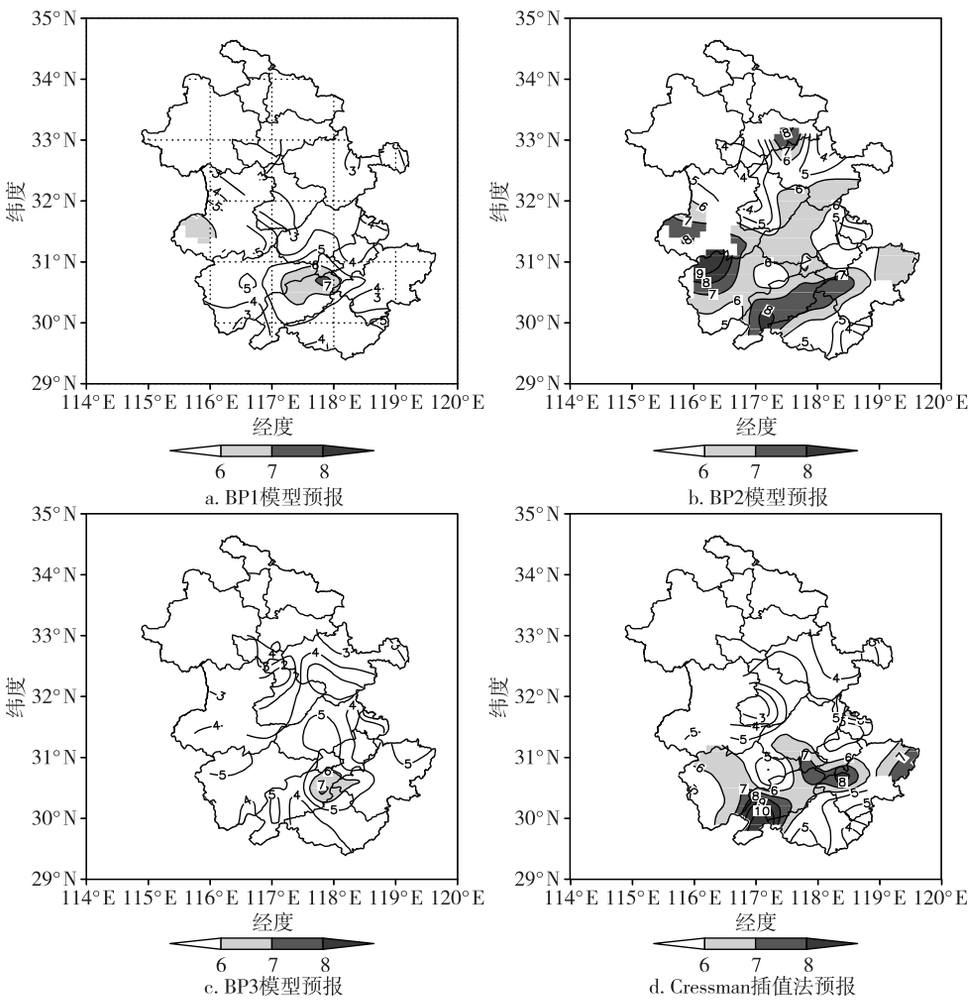


图 4　夏季神经网络降尺度模型及插值方法预报 MAE 空间分布

Fig. 4　The MAE distributions of summer precipitation forecasting results by neural network downscaling schemes and interpolation method, with for BP1, for BP2, for BP3, and for the Cressman interpolation method

表 4　6 种降尺度预报模型分区域指标分析

Table 4　Comparison of forecasting results by 6 downscaling prediction schemes in different regions

区域	模型	BP1 模型	BP2 模型	BP3 模型	LR1 模型	LR2 模型	LR3 模型	插值预报
皖南山区	MAE	4.87	6.23	4.87	6.44	6.26	6.62	5.91
	PR	0.85	0.78	0.87	0.70	0.77	0.69	0.81
沿淮区	MAE	3.36	4.88	3.19	5.35	5.60	5.21	4.82
	PR	0.80	0.74	0.89	0.60	0.62	0.60	0.67
江淮东部区	MAE	3.75	5.98	4.56	5.56	5.91	6.27	5.21
	PR	0.89	0.69	0.79	0.71	0.67	0.65	0.74
大别山区	MAE	5.00	7.31	4.26	6.71	7.09	7.07	5.09
	PR	0.86	0.61	0.90	0.68	0.61	0.63	0.81
沿江西部	MAE	4.64	6.45	4.73	7.23	6.68	7.23	6.59
	PR	0.90	0.79	0.86	0.68	0.71	0.66	0.70
全区	MAE	4.30	6.05	4.43	6.15	6.18	6.40	5.56
	PR	0.86	0.74	0.85	0.68	0.70	0.66	0.76

以看出,BP1 模型和 BP3 模型的预报 MAE 全区平均值分别为 4.30、4.43 mm,PR 分别为 0.86、0.85,而插值预报 MAE 平均值为 5.56 mm,PR 只有 0.76,常用插值预报效果不如这 2 种模型好.分区域来看,BP1 模型在江淮东部和沿江西部预报效果好,BP3 模型在沿淮区与大别山区、皖南山区预报效果较好.无论是分区域还是全区平均来看,预报量和预报因子相同时,以神经网络技术构建模型好于线性回归模型,其中设计的 BP1、BP3 模型好于普通的客观权重插值法,而其余 4 种模型在个别风向上同插值效果相当,但整体没有普通的插值效果好.其原因可能是 BP2 模型中预报因子入选个数较少,其方差贡献率总和较少,及预报量和预报因子之间的关系既有线性又存在非线性,因此预报效果不佳.3 种线性回归模型的预报普遍比插值效果差,可见局地尺度的

降水同大尺度降水场之间以非线性关系为主.综上,BP1 和 BP3 降尺度模型可行,但从网络构建上来看,BP1 模型以单站降水量作为预报值,需建立 46 个神经网络进行训练,构建数目多,需要的训练网络时间较长,而利用 REOF 主成分作为预报量,先对 46 站降水量做降维处理,只需构建 5~15 个神经网络,训练次数减少,计算更加方便.

4 总结与讨论

本文利用中低层风向对夏季资料进行细分,以 3 种预报对象选择单层神经网络和线性回归为降尺度技术,共设计了 6 种统计降尺度模型实现安徽省观测站点逐日降水资料的空间降尺度,并将降尺度预报效果同插值做比对.通过对 MAE、PR 检验指标及降水量总和预报值来进行降尺度拟合和预报效果评估.得出以下结论:

1) 从拟合效果来看,对于相同的预报对象及预报因子而言,神经网络模型在统计评价指标上均好于线性回归模型,可见安徽省局地尺度降水与大尺度降水形势场之间主要是非线性相关的.

2) 通过分析西南风和夏季降水量预报情况,3 种神经网络模型里,以预报对象为单站降水时间序列及局地尺度旋转 EOF 主成分的模型预报效果同插值预报相比有一定的改进,REOF 起到一个降维效果可以有效减少计算工作量,因此具有一定的实际应用价值.BP1 模型在降水极值预报上效果更好,BP3 模型则在 MAE 的标准差、最大值、最小值上表现更为稳定,而预报对象为局地尺度 EOF 主成分的 BP2 降尺度模型效果略差.

3) 分区域来看,BP1 模型在以平原、河流为主要地形的区域预报效果较好,BP3 模型在以山地丘陵地形为主的区域预报效果较好.因此在实际业务运行中,对不同台站应采用不同的预报模型进行预报.

本文为降尺度方法在短期气象资料的应用上拓宽了新的思路,同时也找到新的面向多站点的降尺度方法.随着空间自动加密站资料的进一步完善,利用空间自动加密站资料进行降尺度预报试验是下一步计划.另外在降水预报中会有虚报的情况产生,这时就要结合其他气象要素,建立“消空”指标,使降水预报实现更进一步的精细化预报.

参考文献

References

[1] 刘永和,郭维栋,冯锦明,等.气象资料的统计降尺度方法综述[J].地球科学进展,2011,26(8):837-847
LIU Yonghe,GUO Weidong,FENG Jinming,et al.A summary of methods for statistical downscaling of meteorological data[J]. Advances in Earth Science, 2011, 26(8):837-847

[2] Zorita E, Hughes J P, Lettemaier D P, et al. Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation of local precipitation[J].Journal of Climate,1995,8(5):1023-1042

[3] Charles S P, Bates B C, Hughes J P. A spatiotemporal model for downscaling precipitation occurrence and amounts [J]. Journal of Geophysical Research: Atmospheres,1999,104(D24):31657-31669

[4] 曾晓青,王式功,刘还珠,等.统计降尺度方法在青海省冬季最低温度预测中的应用[J].高原气象,2009,28(6):1471-1477
ZENG Xiaoqing,WANG Shigong,LIU Huanzhu,et al.Application of statistical downscaling in minimum temperature prediction over Qinghai province in winter [J].Plateau Meteorology,2009,28(6):1471-1477

[5] 邱冰,姜加虎,孙占东,等.基于统计降尺度模型的博斯腾湖流域未来气温和降水变化趋势分析[J].资源科学,2010,32(6):1133-1140
QIU Bing,JIANG Jiahu,SUN Zhandong,et al.Analysis of trends in future temperature and precipitation in the Bosten Lake Basin based on a statistical downscaling model[J].Resources Science,2010,32(6):1133-1140

[6] 范丽军,符淙斌,陈德亮.统计降尺度法对华北地区未来区域气温变化情景的预估[J].大气科学,2007,31(5):887-897
FAN Lijun, FU Congbin, CHEN Deliang. estimation of local temperature change scenarios in North China using statistical downscaling method[J].Chinese Journal of Atmospheric Sciences,2007,31(5):887-897

[7] 何慧,金龙,覃志年,等.基于 BP 神经网络模型的广西月降水量降尺度预报[J].热带气象学报,2007,23(1):72-77
HE Hui, JIN Long, QIN Zhinian, et al. Downscaling forecast for the monthly precipitation over Guangxi based on the bp neural network model[J].Journal of Tropical Meteorology,2007,23(1):72-77

[8] 陈丽娟,李维京,张培群,等.降尺度技术在月降水预报中的应用[J].应用气学报,2003,14(6):648-655
CHEN Lijuan,LI Weijing,ZHANG Peiqun,et al.Application of a new downscaling model to monthly precipitation forecast [J]. Quarterly Journal of Applied Meteorology,2003,14(6):648-655

[9] Burrough P A,McDonnell RA.Principles of geographical information system [M]. New York: Oxford University Press,1998

[10] 施能.气象科研与预报中的多元分析方法[M].北京:气象出版社,1995:205-217
SHI Neng. Multivariate analysis methods in weather research and forecasting[M].Beijing:China Meteorological Press,1995:205-217

[11] 金龙.神经网络气象预报建模理论方法与应用[M].北京:气象出版社,2005:41-45
JIN Long.Neural network modeling for weather forecast:

Theories,methods, and applications[M]. Beijing: China Meteorological Press,2005:41-45
[12] 金龙,罗莹,王业宏,等.月降水量的神经网络混合预报模型研究[J].高原气象,2003,22(6):618-623
JIN Long, LUO Ying, WANG Yehong, et al. Study on mixed forecast model of neural network of monthly precipitation[J]. Plateau Meteorology, 2003, 22(6): 618-623

Summer fine precipitation forecasting based on the statistical downscaling technology

HUANG Huirong[1] GUO Pinwen[1]
1 Key Laboratory of Meteorological Disaster of Ministry of Education,
Nanjing University of Information Science & Technology,Nanjing 210044

Abstract Based on the summer daily precipitation data of 60 meteorological stations in Anhui province from 1999 to 2009 and observation data of Anqing sounding station,the relationship between local precipitation and large-scale precipitation circulation in different mid-low wind directions are studied in this paper.The neural network and linear regression method,combined with 3 forecasting objects and corresponding predictor variables are employed to design 6 downscaling function models to approximate and optimize the precipitation data.The 6 models are used to simulate and forecast the daily precipitation data of 46 meteorological stations in Anhui province,and the results show that BP neural network models generally outperform the linear regression models in simulation and forecasting accuracy, indicating the nonlinear correlation between different scales of summer rainfall.Compared with the commonly used Cressman interpolation methods,the neural network models can reflect the basic distribution and local characteristics of summer precipitation in forecasting results.The BP neural network model with single station precipitation series as prediction object has good forecasting results in areas of plains or rivers,while the BP neural network model with the REOF principal components as predicting object is good in mountainous area.

Key words statistical downscaling; daily precipitation; BP neural network model; Rotated Empirical Orthogonal Function(REOF)