

# 基于 PLS 的面雨量统计学降尺度

王靖宇<sup>1</sup> 邱新法<sup>1</sup>

## 摘要

基于 GIS 的面雨量估算方法和基于模式输出的雨量产品都无法解决分辨率过低的问题,并且都不同程度地忽略了中小尺度地形对降水的影响.回顾了各种统计学降尺度方法,使用 NCEP/NCAR 提供的 2011 年 4—9 月的 6 h 一次的再分析资料,以及江苏省气象台提供的全省 20 多个常规站降水实况观测资料,结合高分辨率 DEM 数据,利用偏最小二乘法(PLS)设计了一套考虑地形因子动力作用的面雨量降尺度方案.通过合理选择和构造大尺度预报因子,地形因子动力作用参数化,回归分析与空间插值相结合的面雨量降尺度方案,成功还原了研究区域内代表站的实况降水序列,并绘制出研究区域内高分辨率的面雨量空间分布图.

## 关键词

面雨量;降尺度;地形因子;偏最小二乘

中图分类号 P426.61

文献标志码 A

收稿日期 2012-04-27

资助项目 公益性行业(气象)科研专项(GY-HY200806002)

## 作者简介

王靖宇,男,硕士生,主要研究 GIS 在气象上的应用.542167422@qq.com

邱新法(通信作者),男,博士,教授,博士生导师,主要从事 3S 集成与气象应用方面的研究.xfqiul35@nuist.edu.cn

## 0 引言

通常人们把雨量计在某个特定地点所测得的降落到地面的降水量称作点雨量.为了把雨量数据从空间中的一点转换到有限的空间面积上,面雨量的概念引入气象,使得面雨量的预报成为我国气象学科的一种新的预报对象,是将气象和水文两大应用学科相结合的纽带<sup>[1]</sup>.如何通过观测到的点雨量来准确估算流域的平均降水量,就构成了面雨量估算的核心问题<sup>[2]</sup>.

GIS 因其出色的空间分析能力被充分应用到面雨量的估算中,出现了算数平均法、泰森多边形法、等雨量线法等多种面雨量计算方法.但是基于站点观测数据,利用空间分析的面雨量估算,由于受站点数量及其空间分布的限制,得到的面雨量数据缺乏代表性,只能用于宏观趋势分析.

全球气候模型(GCM)是预测气候变化的主要工具,同样也是降水数据的重要来源.然而,GCM 的空间分辨率往往过低,无法提供尺度小于 200 km 的信息<sup>[3]</sup>.水文过程主要发生在更精细的尺度上,因此,为了更可靠地评估水文过程对气候变化的影响,科研工作者往往需要更高分辨率的气象数据.

为了解决尺度不匹配的问题,在过去的 20 年里,针对获得高分辨率面雨量数据的方法和应用展开了一系列深入研究.基于区域气候模型(RCM)的动力学降尺度方法和基于经验模型的统计学降尺度方法尝试从不同角度解决气候变化情景与影响评估所需分辨率的尺度不一致性.因受限于流参数化、计算量过于繁重等固有缺陷,动力学降尺度无法广泛应用,在这样的背景下,统计学降尺度受到越来越多的关注.

统计学降尺度技术是连接 GCM 低分辨率输出到地表高分辨率气象要素的桥梁<sup>[4-5]</sup>.与动力学降尺度相比,统计学降尺度有以下优点:1)能够以很高的计算效率由大尺度气象要素得到区域尺度的气象要素;2)能输出较高分辨率或站点尺度的气象要素;3)模型参数可以受区域下垫面特征的控制.统计降尺度能够弥补动力降尺度的一些不足,因而也得到了广泛应用.按照所使用原理不同,Wilby 等<sup>[6]</sup>早将统计学降尺度方法划分为随机天气发生器、变换函数法和天气分类法.

随机天气发生器是用于随机生成在统计特征上相似于实际观测

资料的气象数据的一类统计模型,因其能够根据实测数据的统计学特征快速生成高时间分辨率的气象要素序列,随机天气发生器最开始被应用于填补缺失数据,或者替代有问题的历史数据. 应用较为广泛的随机天气发生器有 WGEN<sup>[7]</sup> 和 EARWIG<sup>[8]</sup>. 随机天气发生器由两部分组成,分别描述降水的出现和降水的强度. Gabriel 等<sup>[9]</sup> 提出,使用一阶两态马尔可夫过程对降水的出现进行模拟,按照这种方法,降水序列由有雨日和无雨日构成,降水的发生与否只由前一天的状态决定. 有雨日被判定之后,该日的降水量采用指数分布或 Gamma 分布描述.

变换函数法是应用最早和最为广泛的降尺度技术,通常使用线性和非线性模型建立大尺度气象场与局地气象要素之间的函数关系. 最简单的变换函数是线性模型,大尺度预报因子  $X$  与小尺度预报量  $Y$  的平均值  $\mu$  的关系为

$$\mu = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots, \quad (1)$$

式(1)中  $\beta_i$  表示预报因子  $X_i$  的权重,而预报量  $Y$  则服从均值为  $\mu$  的高斯分布<sup>[10]</sup>. 该类方法主要应用在气候情境资料的降尺度上. 除了简单线性模型和广义线性模型,线性方法还包括典型相关分析(CCA)和奇异值分解(SVD)等. 一般情况下,采用线性降尺度方法就能取得较好的降尺度效果,如 Sailor 等<sup>[11]</sup> 对美国局地气温的模拟, Kidson 等<sup>[12]</sup> 对局地位势高度场的模拟,但是对非线性关系,如极端降水事件的模拟,采用线性降尺度就会产生较大误差.

因为影响降水的因子较多,如果对这些因子进行参数化来进行降尺度建模,过多的因子将增加建模难度,且模型精度没有保障. 而天气分类法则是一种无参数的降尺度方法,它的原理是使用一系列天气类型取代连续的预报因子场,用天气类型  $X_k$  来预报局地降水:

$$Y = f(X_k), \quad (2)$$

式(2)中  $k$  为天气序列的索引,  $Y$  为天气类型  $X_k$  下的局地降水.

天气形势的分类可以依据来源于 GCM 的宏观大尺度气象要素,如海平面气压、位势高度场、气流指数、风向、风速、云量等<sup>[13]</sup>,也可以根据降尺度这一目的对分类准则作出具体改变,建立符合降尺度需求的指标体系. 当某时刻的大尺度天气类别已知时,就可以推测地表区域对应的降水分布以及降水量. 常用的天气分类方法很多,如模糊规则法<sup>[14]</sup>、人工神经网络分类法、气压梯度分类法<sup>[15]</sup>、K 均值分

类法<sup>[16]</sup>、非齐次隐马尔可夫过程等. Wanner 等<sup>[17]</sup> 利用从全球到大陆尺度的变换来重建阿尔卑斯山区的气候变化, Kidson 等<sup>[12]</sup> 运用同种方法对新西兰做了相同的分析.

在以往的面雨量降尺度研究中,不论是动力学降尺度还是统计学降尺度,都仅仅考虑了大尺度气象因子、距离因素等,而忽视了海拔高度和地形因子对降水的影响. 以往的研究表明,海拔数百米的山能使降水增加 50% ~ 80%,更高的山有可能使降水增幅更大. 现在主流模式的水平分辨率一般在 100 km 左右,许多高达数千米的山脉在模式中只能粗略地表示成数百米高的宽阔山脊,这样的分辨率无法真实地刻画山区地形,从而无法准确描述精细格网上发生的地形降水.

根据傅抱璞<sup>[18]</sup> 建立的地形因子和海拔高度对降水的影响模型,坡度、坡向、海拔高度对降水的增幅起到非常重要的作用,因此本文考虑坡度、坡向以及海拔高度建立面雨量降尺度模型.

## 1 资料来源与数据处理

### 1.1 地形因子的提取

DEM 因其能够在最大程度上接近地面真实信息,所以普遍被各国政府规定为敏感数据,较难获得,且不能正式对外发布. 因此必须选择使用接近 DEM 的其他数据源,本文采用 SRTM3 (Ver 4.1) 这一精度较高的免费数据集作为研究区域地形信息的数据来源.

从坡度、坡向的概念出发,根据李占清等<sup>[19]</sup> 提出的计算模型,计算得到的江苏省 5 km 分辨率的海拔高度、坡度和坡向,如图 1 所示.

### 1.2 大尺度预报因子的选择

在建统计学降尺度模型的过程中,预报因子的选择非常重要. 判断是否选择某个预报因子的最基本依据是,它是否包含有用信息,也就是它是否具有很强的预报能力. 合适的预报因子可以通过统计学分析,以及其与预报量之间的物理过程相关性进行遴选. 湿度、温度等表征大气环流特征的预报因子通常被应用到降水的降尺度研究中, Charles 等<sup>[5]</sup> 提出,对相对湿度的测量,比如温度露点差等,较比湿的测量在降水的降尺度研究中更为重要.

一般来说,大尺度预报因子的选择遵从以下 4 个标准<sup>[20]</sup>:

1) 选择的预报因子要与所预报的预报量有很

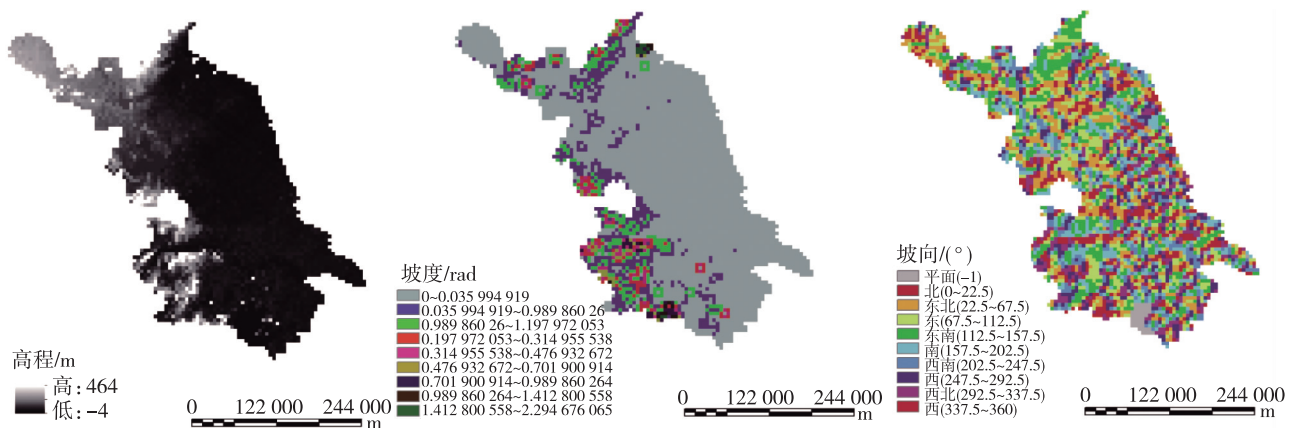


图1 江苏省5 km分辨率海拔高度、坡度、坡向

Fig. 1 The elevation, slope and aspect chart of Jiangsu province with 5 km resolution

强的相关;

2) 预报因子能够代表大尺度气象要素的重要物理过程以及变化性;

3) 能够被 GCM 精确模拟,从而减小甚至消除来源于 GCM 的系统误差;

4) 所选择的预报因子之间应该是若相关或无关的。

本文采用 NCEP 再分析资料作为大尺度预报因子数据源,但是由于再分析资料提供的数据产品种类繁多,如果将所有地面以及 17 个等压面气象要素数据全部选作预报因子,势必会在极大地增加模型的复杂程度,降低模型运转效率,根据相关降尺度研究的经验,预报因子的选取一般为 6 ~ 12 个左右为佳。

对于高层大气的预报因子,本文最终选定 700 和 850 hPa 2 层水汽通量散度平均值,700 和 850 hPa 2 层水平方向全水汽通量平均值,500、600 和 700 hPa 3 层相对湿度平均值,500、600 和 700 hPa 3 层露点温度差,共 4 个变量。

对于地面的预报因子,本文选定海平面气压、相对湿度和可降水量。

为了检验所选出的大尺度预报因子对降水的预

报能力,本文选用 4 个代表站(南京、南通、盐城、淮安)自 2010 年 4 月 1 日 00 时—2010 年 9 月 30 日 18 时,间隔为 6 h 的实测降水序列,分别对 7 个预报因子对应的时间序列进行相关性检验,表 1 给出了入选因子与预报对象之间的相关系数。

最终确定的预报因子集,其中既包含基本因子,又包含组合因子,既有动力学因子,也有热力学因子,普遍具有丰富的天气学意义。因子与降水的相关系数较高,相关系数大多在 0.3 以上,最高的超过 0.5,对降水具有较强的预报能力。

为了体现大尺度预报因子对应格网内各个次格网的差异,同时综合考虑地形因子和海拔高度对宏观大气背景的影响,本文分别将各次格网的海拔高度、坡度、坡向作为第 2、3、4 影响因素引入到各个大尺度预报因子的空间插值中,获得了考虑地形因子的高空间分辨率大尺度预报因子数据集。

## 2 降尺度模型的建立

### 2.1 预报对象的处理

江苏省全境共分布 74 个基础气象站,尤其在南京、盐城、南通等地分布较为密集。如果分别对 74 个基础站分别建立模型,工作量过大,因此本文采用分

表 1 预报因子与各代表站降水量的相关系数

Table 1 Correlation coefficient between predictors and precipitation amount at each representative station

站点	平均全水汽通量	平均水汽通量散度	平均相对湿度	平均露点温度差	大气可降水量	地面相对湿度	海平面气压
南京(58238)	0.532	-0.454	0.371	0.335	0.356	0.318	-0.351
南通(58259)	0.517	-0.478	0.398	0.344	0.356	0.299	-0.336
盐城(58154)	0.493	-0.466	0.385	0.34	0.375	0.267	-0.314
淮安(58141)	0.497	-0.473	0.383	0.335	0.344	0.301	-0.331

组建模的方式,将整个研究区域划分为若干子区域,在每个子区域中分别选择代表站进行面雨量降尺度建模.

根据大尺度气象要素格网覆盖情况,得出部分站点分组方案如表 2.

表 2 部分站点分组方案

Table 2 Meteorological station grouping scheme

区号	代表站	相关站
I	南京(58238)	58235,58237
II	南通(58259)	58252,28253,58254
III	盐城(58154)	58150,58158
IV	淮安(58141)	58132,58139

2.2 建模原理

由于子研究区域的大尺度预报因子条件相同,区域内除代表站外的其他次格网的差异性主要集中在地形因子和海拔高度上,而这种差异进而通过表征该区域的降尺度模型,体现为各个次网格降水量的不同.模型的实质在于解释不同次网格的海拔高度和地形因子对相同大尺度气象条件的响应机制.

建模原理见图 2,图 2 中  $X_i$  代表大尺度气象因子,红色的  $P$  代表参与建模的代表站所处格网的降水量,蓝色的  $P$  代表研究区域内其他格网的降水量,二者的区别体现在海拔高度  $h$ 、坡度  $\alpha$ 、坡向  $\beta$  的不同.

2.3 代表站模型的建立

回归模型作为降尺度研究中应用最为广泛的模型,长久以来受到广泛的认可.回归建模的目的在于通过分析观测数据建立统计学模型,辨识数据中的系统因素,并将其与随机误差区别开,达到  $Y = BX +$

$\epsilon$  的建模目的( $Y$ 和 $X$ 分别是因变量和自变量向量, $B$ 为系数向量, $\epsilon$ 为随机误差向量).

因此,本文拟采用回归分析法,建立的如下面雨量降尺度模型:

$$P = a_0 + a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4 + a_5X_5 + a_6X_6 + a_7X_7 + \epsilon, \quad (3)$$

其中, $X_i$ 为大尺度预报因子的考虑地形因子及海拔高度的协克里格插值结果.

常用的多元统计回归法在估算降水的时空分布时,计算值与实测值存在较大的误差<sup>[21]</sup>.究其原因不外乎两个:1)自变量系统对预报量的解释性不足;2)自变量之间的多重相关性使得某些重要影响因子的权重被低估.

偏最小二乘(PLS)方法吸取了主成分分析的精髓,采取先提取成分的方式,首先从自变量矩阵  $X$  中提取成分向量  $t_1$ ,它既是各自变量的线性组合又能最大程度的携带  $X$  中的数据变异.假定因变量  $Y = Bt_1$ ,再基于  $t_1$  对  $Y$  具有最大解释能力,即二者相关性最高的约束条件下进行  $Y$  对  $t_1$  的回归,得到  $Y = B_{PLS1}t_1 + \epsilon_Y$ ,然后进行交叉有效性检验,如果模型没有达到满意的精度,则利用  $X$  和  $Y$  被  $t_1$  解释后残差  $\epsilon_X$  和  $\epsilon_Y$  进行第 2 轮成分提取,直到达到满意的精度.最终,若通过交叉有效性检验的成分个数为  $m$ ,则将因变量  $Y = \sum_{i=1}^m B_{PLSi}t_i$  还原成关于原始变量的回归方程  $Y = B_{PLS}X + \epsilon_{PLS}$ , PLS 回归建模完成<sup>[22]</sup>.

根据原理,本文参考文献[23-28]发布的 PLS 算法,建立了 4 个研究区域 00、06、12 和 18 时以及全天的降水量的空间分布模型(表 3).由于降水时空分布的最大特点是不连续性,模型的主旨在于对

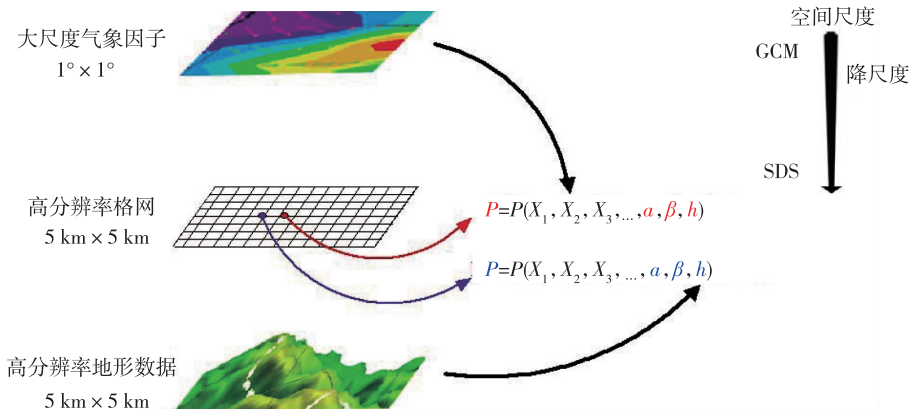


图 2 降尺度建模原理

Fig. 2 Conceptual chart of statistical downscaling

表3 分时段降尺度模型

Table 3 6-hour and daily downscaling models

区号	时次	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	$a_7$
I	00 时	-0.17	0.07	0.34	1.48	-3.42	0.20	-0.01
	06 时	-0.51	0.19	0.86	0.00	13.17	0.14	0.06
	12 时	-0.88	0.64	0.71	0.07	10.11	-0.06	0.12
	18 时	-0.01	0.17	-0.11	-0.15	-1.44	0.25	0.00
	日	-1.41	0.85	1.74	0.07	9.12	0.10	0.09
II	00 时	-0.36	0.11	0.53	1.29	5.29	-0.31	0.01
	06 时	-0.19	0.12	0.14	0.11	-8.39	0.09	0.04
	12 时	-0.43	1.11	1.23	0.12	17.48	0.10	0.21
	18 时	-0.15	0.07	0.04	-0.05	0.52	0.09	0.04
	日	-1.03	0.62	1.27	0.05	10.54	0.07	0.06
III	日	-0.79	1.12	2.21	0.16	7.90	0.17	0.14
IV	日	-1.98	1.26	1.48	0.08	8.89	0.09	0.15

面雨量的降尺度,而非对降水概率的降尺度,如果对整个时间序列建立模型必将导致模型对无雨时次做出有雨的判定,并且直接引入常数项  $a_0$  也会造成模型对无雨区输出降水量,因此本文拟采用大尺度气象要素集中的降水率序列替代常数项  $a_0$ ,针对实测降水量与大尺度降水率的差异建立 PLS 回归模型,这样做的优点是能够更集中讨论地形因子以及海拔高度对区域气象条件的响应,避免误判。

### 3 降尺度结果分析

为了检验降尺度模型在未来气候条件下的适用性,需要对降尺度模型的精度进行检验。比较经典的检验方法为历史数据检测法。选定一段可获得的历史观测数据,使用待检验的降尺度方案对其处理,然后通过对比其降尺度结果与独立记录的观测值,来进行验证所使用降尺度模型的精度。当具备很长的历史观测记录序列,并且该序列包含较大变化率时,这种检验方法较为适用。

但是对历史数据的准确拟合并不能证明模型在复杂气候条件下的普遍适用性,因此本文采用代表站所在组内其他站点的模型输出结果与实况数据,对降尺度模型进行适用性检验。

#### 3.1 代表站检验

##### 3.1.1 降尺度结果与实况数据趋势对比

通过对各代表站的实况数据与降尺度结果进行对比,可以较为直观地了解本文所使用降尺度方案的整体效果。由代表站 58238(图 3)各时次对比结果可以看出,采用 PLS 方法建立的降尺度模型能够较

好地还原代表站的主要降水过程,对适度降水拟合效果较好,但是对弱降水( $< 5 \text{ mm/d}$ )和强降水( $> 40 \text{ mm/d}$ )的模拟较差,如对 7 月 14 日和 16 日的强降水,降尺度结果较实况数据的误差接近 50%。总体来说,所建立的面雨量降尺度模型仍有高估弱降水、低估强降水的趋势。

##### 3.1.2 降尺度结果整体精度评价

为了检验面雨量降尺度模型的精度,通常要对综合考虑降水量的平均值、变化性以及分位点<sup>[29-30]</sup>。为了将降尺度结果序列与实况降水序列各自的标准差,以及二者之间的相关系数进行可视化展现,本文采用泰勒图表<sup>[31]</sup>对降尺度模型的精度进行整体评价。

图 4、5 中,坐标轴长度为降尺度结果序列的标准差与实况数据序列的标准差之间的比值,角度为降尺度结果序列与实况数据序列的相关系数。

观察对比图 4 和图 5 可以得出这样的结论:通过之前建立的面雨量降尺度模型得到的各个代表站降尺度结果序列与实况降水数据之间的相关性较好,相关系数在 0.80 与 0.92 之间,说明模型很好地还原了降水的时间序列以及总体趋势,但是各代表站的降尺度结果标准差普遍小于对应的实况降水标准差,说明模型虽然能够在一定程度上体现降水强度的变化性,可这种变化性无法在模型中得到显著表达,也就是说,模型对降水强度的变化性起到了平滑的效果。

#### 3.2 适用性检验

##### 3.2.1 检验标准

为了定量检验降尺度方案的精度,本文根据实

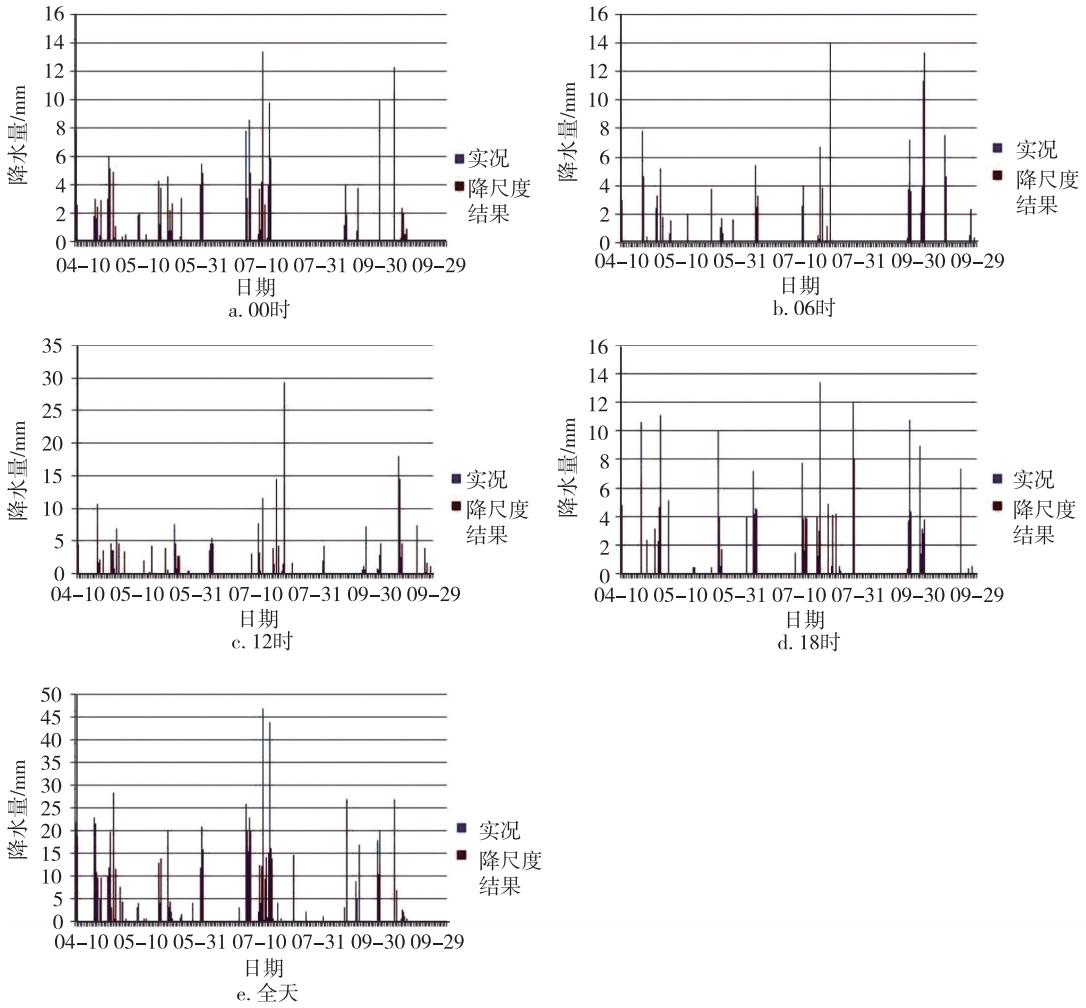


图3 2010年4—9月58238站各时次降水实况降尺度结果对比

Fig. 3 Comparison between observation and downscaled outcomes at station 58238 from April to September, 2010

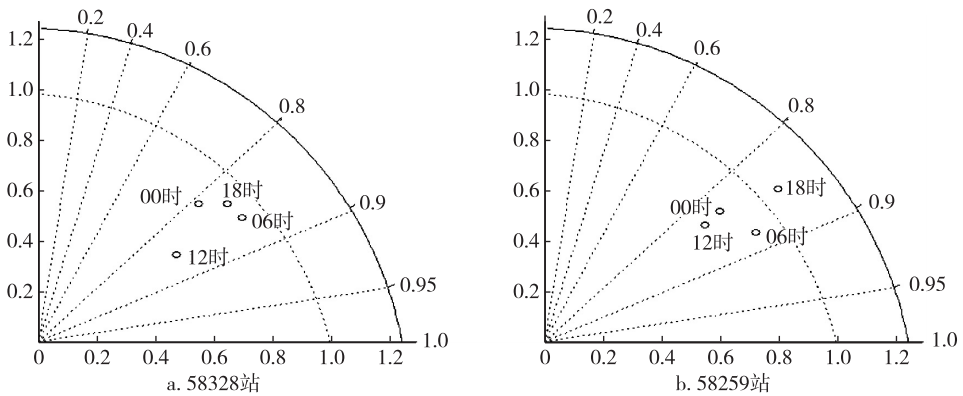


图4 58238站和58259站分时段泰勒图

Fig. 4 6-hour Taylor diagram of station 58238 and 58259

际业务中常用的单站降水和温度预报检验方法<sup>[32-34]</sup>,设计了平均绝对误差和预报准确率2个指标,分别定义如下:

$$T_{MAE} = \frac{1}{N} \sum_{i=1}^N |P_i - O_i|, \tag{4}$$

$$T_T = \frac{N_p}{N_f} \times 100\%, \quad (5)$$

式中,  $P_i$  为待检验站  $i$  时次的降尺度结果,  $O_i$  为该站  $i$  时次的实况数据,  $N_p$  代表  $|P_i - O_i| \leq 20\%$  的次数,  $N_f$  为降水发生的总时次, 准确率  $T_T$  的含义是面雨量降尺度结果与实测数据误差不超过实测数据 20% 的百分率。

### 3.2.2 检验分析

各代表站及所在组内其他站点的平均绝对误差和准确率如表 4 及 5 所示。

通过观察不难发现, 降尺度模型对所有站点不同时次的降尺度效果都有一个共同的特点, 即 00 时的平均绝对误差最低, 相应的预报准确率也最高, 对 12 时的效果最差。通过分析不难得知, 午间往往对应一日之中最强的对流活动, 一日中的降水极值通常发生在这一时次, 大尺度预报因子也在此区间产生较大的变化率, 而这些短时极端变化在 6 h 时间分辨率中很难得到体现。这也很好地解释了为什么降尺度模型对降水极值的模拟具有明显的滞后性和平均性, 而对其他各时次以及全天的总体模拟效果优于 12 时。因此, 本文提出的针对 12 时的面雨量降

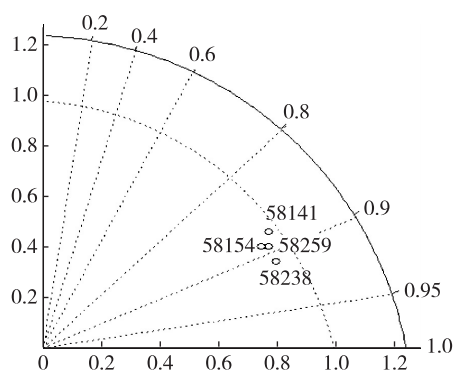


图 5 各代表站逐日泰勒图表

Fig. 5 Daily Taylor diagram of each relevant station

尺度模型具有一定的参考价值, 对其他时次和全天的模型均具有较好的实用性。

### 3.3 研究区降水空间分布的拟合

基于通过代表站检验、适用性检验的面雨量降尺度模型, 本文绘制出了空间分辨率为 5 km 的 4 个子研究区域在 2010 年 7 月 12 日波及江苏省绝大多数地区的一次强降水过程中的降水量空间分布 (图 6)。

表 4 代表站各时次降水量的平均绝对误差与准确率

Table 4 Average absolute errors and accuracy of each representative station at different time

站点	平均绝对误差/mm					准确率/%				
	00 时	06 时	12 时	18 时	日	00 时	06 时	12 时	18 时	日
58238	2.79	2.81	6.28	2.29	3.28	81	74	53	75	68
58259	2.91	3.09	5.25	2.25	4.50	78	70	50	76	72
58154	N/A	N/A	N/A	N/A	3.93	N/A	N/A	N/A	N/A	73
58141	N/A	N/A	N/A	N/A	5.41	N/A	N/A	N/A	N/A	75

表 5 相关站各时次降水量的平均绝对误差与准确率

Table 5 Average absolute errors and accuracy of each relevant station at different time

站点	平均绝对误差/mm					准确率/%				
	00 时	06 时	12 时	18 时	日	00 时	06 时	12 时	18 时	日
58235	3.57	2.09	4.79	3.4	5.57	68	70	49	79	67
58237	2.41	5.75	6.71	3.36	5.74	65	68	42	76	61
58252	1.79	2.72	5.11	2.72	4.35	82	71	52	65	69
58253	2.46	4.53	7.20	1.10	6.53	70	67	54	66	68
58254	1.51	3.58	4.35	3.27	4.17	77	72	61	59	57
58150	N/A	N/A	N/A	N/A	6.13	N/A	N/A	N/A	N/A	59
58158	N/A	N/A	N/A	N/A	5.01	N/A	N/A	N/A	N/A	63
58132	N/A	N/A	N/A	N/A	5.19	N/A	N/A	N/A	N/A	59
58139	N/A	N/A	N/A	N/A	4.78	N/A	N/A	N/A	N/A	67

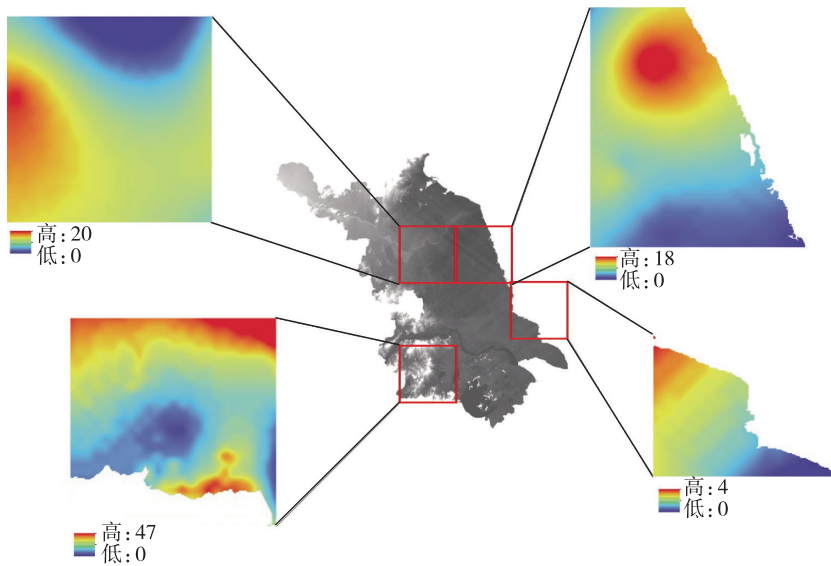


图6 2010年7月12日一次强降水过程中各子研究区域降水量空间分布(5 km)

Fig.6 Precipitation distribution in sub-survey regions during a heavy rain at July 12, 2010

## 4 结论

本文基于 NCEP/NCAR 提供的 2010 年 4—9 月的空间分辨率为  $1^\circ \times 1^\circ$  时间分辨率为 6 h 的再分析资料,以及江苏省气象台提供的相应时段全省 20 余个气象站间隔 1 h 的实况降水资料,结合由江苏省高空间分辨率 DEM 数据提取的地形因子和海拔高度等数据,以偏最小二乘回归分析为基础,结合协克里格插值等统计学方法,选取江苏省内 4 个代表区域进行面雨量降尺度建模,模拟了区域内各站点 00、06、12 和 18 时以及逐日的降水状况.在整个试验过程中,获得以下主要结果:

1) 相对传统降尺度方法选择只具有单一意义的基本因子,如某层等压面上的经向风速、相对湿度等,对基本因子进行必要的组合和变换,从而获得对降水更具明确物理意义的复合因子,如若干层的平均水汽通量散度、全水汽通量等,能够提高预报因子与降水量的物理过程相关性,增强预报因子所包含的信息量,降低各预报因子之间的信息冗余,从而进一步保证面雨量降尺度模型的物理过程连续性,满足终端用户对面雨量降尺度的要求,使模型能够更好地解释大尺度气象要素影响降水的物理机理.

2) 地形因子对降水量有着显著影响,地形因子的参数化对于提高降尺度模型的精度以及解释真实发生在次网格中的动力学过程,如辐射过程、对流过程、云微物理过程等有着重要的意义,并且地形因子参数化能够更好地体现不同次格网之间的个体差

异,提高信噪比,使得降尺度模型能够更好地剖析不同地区的地形因子对相同气象背景的响应机制,进而改善降尺度模型的模拟精度.

3) 根据大尺度预报因子的格网将研究区域划分为若干子区域,从每个子区域众多站点中合理地选取代表站建立面雨量降尺度模型的思路是可行的.经过宏观趋势对比,整体精度评价,以及相关站的平均误差检验和准确率检验,基于偏最小二乘回归法(PLS)建立的面雨量降尺度模型能够较好地还原发生在研究区域内的主要降水过程.

4) 总体来看,通过合理选择大尺度预报因子,并对其重新组合和变换以获得与降水量相关程度更高的预报因子,地形因子动力作用参数化,采用回归分析和空间插值相结合的面雨量降尺度方案具有较高的实际应用价值,并且所建立的面雨量降尺度模型可以直接使用数值预报产品作为模型输入量,因此具有操作简便、使用方便的优点,便于在基层台站推广使用.

## 参考文献

### References

- [1] 董官臣,冷林茂,符长锋. 面雨量在天气预报中的应用[J]. 气象, 2000, 26(1): 9-13  
DONG Guanchen, YE Linmao, FU Changfeng. An application of area rainfall rainfall to weather forecast[J]. Meteorological Monthly, 2000, 26(1): 9-13
- [2] 徐晶,姚学祥. 流域面雨量估算技术综述[J]. 气象, 2007, 33(7): 15-21



- XU Jing, YAO Xuexiang. Watershed areal precipitation estimation technology: A review [J]. *Meteorological Monthly*, 2007, 33(7):15-21
- [3] Intergovernmental Panel on Climate Change. *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Inter-governmental Panel on Climate Change, Freshwater resources and their management* [M]. Cambridge: Cambridge University Press, 2007
- [4] Wilby R L, Dawson C W, Barrow E M. SDSM: A decision support tool for the assessment of regional climate change impacts [J]. *Environmental Modelling & Software*, 2002, 17(2):147-159
- [5] Charles S P, Bates B C, Hughes J P. A spatiotemporal model for downscaling precipitation occurrence and amounts [J]. *Journal of Geophysical Research*, 1999, 104(D24):31657-31669
- [6] Wilby R L, Wigley T M L. Downscaling general circulation model output: A review of methods and limitations [J]. *Progress in Physical Geography*, 1997, 21(4):530-548
- [7] Richardson C W. Stochastic simulation of daily precipitation, temperature, and solar radiation [J]. *Wat Resour Res*, 1981, 17(1):81-82
- [8] Kilsby C G, Jones P D, Burton A, et al. A daily weather generator for use in climate change studies [J]. *Environ Modell Softw*, 2007, 22(12):1705-1719
- [9] Gabriel K R, Neumann J. A Markov chain model for daily rainfall in occurrence at Tel Aviv [J]. *Q J R Meteorol Soc*, 1962, 88(375):90-95
- [10] 杨赤, 严中伟, 邵月红. 基于 Tweedie 分布的日降水量统计降尺度模型 [J]. *北京师范大学学报:自然科学版*, 2009, 45(5):531-535  
YANG Chi, YAN Zhongwei, SHAO Yuehong. Statistical downscale model for daily precipitation based on Tweedie distribution [J]. *Journal of Beijing Normal University: Natural Science*, 2009, 45(5):531-535
- [11] Sailor D J, Li X. A semiempirical downscaling approach for predicting regional temperature impacts associated with climatic change [J]. *J Climate*, 1999, 12(1):103-114
- [12] Kidson J W, Thompson C S. A comparison of statistical and model-based downscaling techniques for estimating local climate variations [J]. *J Climate*, 1998, 11(4):735-753
- [13] Conway D, Wilby R L, Jones P D. Precipitation and air flow indices over the British Isles [J]. *Climate Research*, 1996, 7(2):169-183
- [14] Bardossy A, Stehlik J, Caspary H J. Automated objective classification of daily circulation patterns for precipitation and temperature downscaling based on optimized fuzzy rules [J]. *Climate Research*, 2002, 23(1):11-22
- [15] Ghosh S, Mujumdar P P. Future rainfall scenario over Orissa with GCM projections by statistical downscaling [J]. *Current Science*, 2006, 90(3):396-404
- [16] Wetterhall F, Halldin S, Xu C Y. Seasonality properties of four statistical-downscaling methods in central Sweden [J]. *Theoretical and Applied Climatology*, 2007, 87(1):123-137
- [17] Wanner H, Rickli R, Salvisberg E, et al. Global climate change and variability and its influence on Alpine climate: Concepts and observations [J]. *Theoretical and Applied Climatology*, 1997, 58(3/4):221-243
- [18] 傅抱璞. 地形和海拔高度对降水的影响 [J]. *地理学报*, 1992, 47(4):302-314  
FO Baopu. The effect of topography and elevation on precipitation [J]. *Acta Geographica Sinica*, 1992, 47(4):302-312
- [19] 李占清, 翁笃鸣. 一个计算山地地形参数的计算机模式 [J]. *地理学报*, 1987, 42(3):269-274  
LI Zhanqing, WENG Duming. A computer model to determine topographic parameters [J]. *Acta Geographica Sinica*, 1987, 42(3):269-274
- [20] 范丽军, 符淙斌, 陈德亮. 统计降尺度法对未来区域气候变化情景预估的研究进展 [J]. *地球科学进展*, 2005, 20(3):320-329  
FAN Lijun, FU Congbin, CHEN Deliang. Review on creating future climate changes scenarios by statistical downscaling techniques [J]. *Advance in Earth Science*, 2005, 20(3):320-329
- [21] 周锁铨, 薛根元, 周丽峰, 等. 基于 GIS 降水空间分析的逐步插值方法 [J]. *气象学报*, 2006, 64(1):100-111  
ZHOU Suoquan, XUE Genyuan, ZHOU Lifeng, et al. The stepwise interpolation approach of precipitation for spatial analysis based on GIS [J]. *Acta Meteorologica Sinica*, 2006, 64(1):100-111
- [22] Partial Least Squares (PLS) [EB/OL] (2011-10-10). <http://www.statsoft.com/textbook/partial-least-squares/>
- [23] Martens H, Naes T. *Multivariate calibration* [M]. Chichester: John Wiley & Sons, 1989
- [24] Höskuldsson A. PLS regression methods [J]. *Journal of Chemometrics*, 1988, 2(3):211-228
- [25] Eriksson L, Johansson E, Kettaneh-Wold N, et al. *Multi- and megavariate data analysis: Principles and applications* [M]. Umeå: Umetrics, 2001
- [26] Palyulin V A, Radchenko E V, Baranova O D, et al. MFTA: Recent extensions of molecular field topology analysis, in EuroQSAR2002 designing drugs and crop protectants: Processes, problems and solutions [M]. Bournemouth: Blackwell Publishing, 2003:188-190
- [27] Palyulin V A, Radchenko E V, Zefirov N S. Molecular field topology analysis method in QSAR studies of organic compounds [J]. *J Chem Inf Comput Sci*, 2000, 40(3):659-667
- [28] Frei C, Scholl R, Fukutome S, et al. Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models [J]. *J Geophys Res*, 2006, 111, D06, 105, doi:10.1029/2005JD005965
- [29] Halenka T, Kalvov J, Chldov Z, et al. On the capability of RegCM to capture extremes in long term regional climate simulation-comparison with the observations for Czech Republic [J]. *Theoretical and Applied Climatology*, 2006, 86(1/2/3/4):125-145
- [30] Maraun D, Rust H W, Osborn T J. Synoptic air flow and UK daily precipitation extremes: Development and validation of a vector generalised linear model [J]. *Extremes*,

- 2010,13(2):133-153
- [31] Frei C, Chistensen J H, Deque M, et al. Daily precipitation statistics in regional climate models; Evaluation and intercomparison for the European Alps [J]. *J Geophys Res*, 2003, 108(D3):4124-4128
- [32] 中国气象局科教司. 省地气象台短期预报岗位培训教材[M]. 北京:气象出版社, 1998:25-40  
Department of Science and Education, China Meteorological Administration. Short-term forecasting training materials for provincial meteorological stations [M]. Beijing: China Meteorological Press, 1998:25-40
- [33] 黄海洪, 郑凤琴, 孙崇智. 一种新型降水预报评分方法[J]. *气象*, 2004, 30(9):35-38  
HUANG Haihong, ZHENG Fengqin, SUN Chongzhi. Application study of new evaluation method for rainfall forecast [J]. *Meteorological Monthly*, 2004, 30(9):35-38
- [34] 中国气象局. 中短期天气预报质量检验办法(试行)[M]. 北京, 2005  
China Meteorological Administration. The evaluation method for short-term weather forecast [M]. Beijing, 2005

## Research on downscaling in areal precipitation estimation based on PLS regression

WANG Jingyu<sup>1</sup> QIU Xinfu<sup>1</sup>

<sup>1</sup> School of Remote Sensing, Nanjing University of Information Science & Technology, Nanjing 210044

**Abstract** Both GIS-based areal precipitation estimation and gridded precipitation data from GCM outputs have the problem of poor resolution, and to some extent neglect the influence of small or medium landform on precipitation distribution. A variety of statistical downscaling methods are briefly reviewed in this paper. Taking into account of the dynamic effect of geographic factors, a new downscaling scheme for areal precipitation estimation is proposed using 6 hour temporal resolution re-analysis data of NCEP/NCAR from April to September of 2010, and observational precipitation data from more than 20 weather stations in Jiangsu. The scheme chooses and constructs appropriate large-scale predictor, and retrieves the small-scale geographic factors from high resolution DEM, parameterizes the dynamic effect of geographic factors, and integrates the regression analysis, Partial Least Squares (PLS) and spatial interpolation. The actual precipitation series in the researched weather stations are successfully retrieved by the proposed scheme and the high-resolution spatial distribution of areal precipitation are drawn.

**Key words** areal precipitation; downscaling; geographic factors; Partial Least Squares (PLS)