

客观定量预报中降水的预处理^{* 1}

赵声蓉¹⁾ 裴海瑛²⁾

¹⁾(国家气象中心,北京 100081) ²⁾(江苏省气象台,南京 210008)

摘 要

建立预报模型前,对降水量进行一定的处理会对预报效果有较大的影响。对于降水量为 0 的样本,根据对应的相对湿度情况分别赋予 0 或不同的负值,并通过神经网络方法,以中国国家气象中心 T213 模式、德国气象局业务模式和日本气象厅业务模式相应的降水量预报结果作为预报因子,利用 2003 年和 2004 年夏季资料分别建立了处理后降水量以及未经处理降水量的预报模型。以北京等站为例,2005 年夏季试报结果的对比分析表明:通过相对湿度对降水量进行适当处理后,预报结果从 TS 评分、空报率、漏报率及预报偏差来说,不论是与不进行处理的预报结果还是与模式直接的预报结果相比都有提高,尤其是减少了空报的情况。该处理方法简单可行,并且对降水预报效果提高明显。

关键词: 降水预处理; 空报率; 漏报率; 预报偏差

引 言

随着数值预报技术的发展,基于数值预报产品释用的客观要素预报已成为日常预报业务中的重要参考手段,同时预报业务的需要也大大促进了数值预报产品释用技术的研究和应用。我国气象工作者在这方面进行了大量的工作,很多的研究成果在实际预报业务中起到了重要的作用^[1-4]。

客观降水预报一直是数值预报产品释用中的难点问题,相比较温度和湿度预报来说,降水预报效果一直不是很理想。近年来,我国气象工作者就如何提高客观降水预报效果进行了大量的研究工作,这些工作大体分为两个方面,一是从预报因子上下功夫,寻找更合理的预报因子,比如在预报因子中加入了数值化的预报经验^[5-6];另一方面是在建立预报模型前对降水进行一定的预处理^[7-8],对降水量的处理主要是分级处理,或者分别建立不同量级降水概率预报模型,这些研究工作在一定程度上取得了提高预报效果的作用。

客观降水预报效果较差,尤其是客观定量降水预报效果较差的重要原因是降水不是连续变量,也

不是一个时间点上的观测值,而是一段时间内的累积值,同时降水是严重偏态分布的变量,利用传统的统计方法直接对降水量进行预报是存在问题的,因此人们往往采取分级的方法对降水量进行处理^[9]。但降水还存在与其他要素完全不同的特点:不同的天气形势在降水量中没有得到完全的反映,换句话说,对于零降水量的情况包含了很多种不同的信息,例如,雷声很大但没有下雨,湿度很大也没有下雨,天空有厚厚的层云也没有下雨,天空没有一丝云同样没有下雨,这些天气形势千差万别,但在降水量中却得不到反映,这使得在建立预报方程时因子与预报量的关系不能得到很好的反映,影响了预报模型的质量。这个问题在分级处理中是得不到解决的,因此在建立预报模型前,如果通过一定的方式对降水量进行处理,使得不同零降水量的情况在处理后的序列中得到一定程度的区分,而不是直接用降水量建立预报模型可能会对预报结果产生较大的影响。

另一方面,从预报业务的需求来说,分级降水预报不能完全满足需要,与气象要素有关的一些专业预报,比如,地质灾害预报、森林火险预报等需要的是定量降水预报结果,因此客观定量降水预报问题

* 中国气象局 2005 年气象新技术推广项目“精细化大气要素中短期客观预报业务系统开发”(CMATG2005M07)和国家科技攻关计划奥运科技专项项目“奥运气象保障技术研究”(2002BA904B05)共同资助。

2005-09-13 收到,2006-06-26 收到再改稿。

也就变得重要起来。

基于以上的考虑,为了探讨提高定量降水预报效果的有效途径,本文将利用相对湿度来对降水量进行预处理,并就同样方法、同样因子、同样时段、同样参数,分别建立未处理和处理后的降水量预报模型,进行试报,通过以下的分析可以看到对降水量的这种处理方式能够提高预报效果。

1 神经网络方法简介

文中采用神经网络中的 BP 网络方法,它是一种单向传播的多层前向网络^[10]。BP 网络包括了输入层、隐层和输出层 3 个部分,隐层可以是多层也可以是一层,隐层上的节点称为隐节点。输入信号向前传播到隐层节点,通过权值和作用函数的作用形成隐节点的值,隐节点的值再向前传播到下一层,最后获得输出结果,每一层节点的输出只影响到下一层节点的输出。作用函数决定了网络节点的特性,设 o_i 为节点 i 的输出, n_i 为节点 i 的输入,则 $o_i = f(n_i)$,其中 $f(x)$ 为作用函数, $n_i = \sum_j w_{ij} o_j$, w_{ij} 为节点 i 对应上一层节点 j 的权值, o_j 为上一层节点 j 对应节点 i 的输出。

BP 网络可以看成是一个从输入到输出的高度非线性映射。作用函数通常为 Sigmoid 型函数,本文中采用的作用函数如下所示:

$$f(x) = \frac{1 - e^{-x}}{1 + e^{-x}}$$

网络的权值通过学习获得,这里的学习方法采用的是误差反向传播算法,误差反向传播算法的基本过程为:① 初始化,即给网络权值赋以初值,一般是赋予一个不为零但接近零的值;② 有了网络权值以后分别计算各个样本隐层和输出层的值,并由此可以得到输出结果与实况之间的误差;③ 根据误差结果由后向前采用梯度下降法修正各层各节点的权值;④ 重复②,③直到网络收敛或达到一定的误差要求,学习结束。

2 资料处理及试验方案设计

2.1 资料及处理方法

本文所使用的资料为中国国家气象中心 T213 模式(以下简称 T213 模式)、德国气象局业务模式(以下简称德国模式)和日本气象厅业务模式(以下简称日本模式)12:00(世界时,下同)起报的降水量

预报以及我国多个站点一天 8 次的观测资料。

T213 模式资料为 $0.5625^\circ \times 0.5625^\circ$ 的经纬网格点资料;德国模式资料为 $1.5^\circ \times 1.5^\circ$ 的经纬网格点资料;日本模式资料为 $1.25^\circ \times 1.25^\circ$ 的经纬网格点资料。

预报对象为 00:00 的 24 h 累积降水量,预报因子为 3 个模式不同预报时效与实况相应的 24 h 累积降水量预报,即 12~36 h 的累积降水量预报,36~60 h 的累积降水量预报,依此类推。为了避免由插值引起的误差,因子不插值到站点,而以站点周围 4 个点都作为预报因子。

对降水量的预处理过程如下:① 由观测资料计算降水时段内的平均相对湿度;② 计算建模时段内逐日的平均相对湿度标准化变量 f_{hs} ;③ 计算建模时段内平均降水量 e_r ;④ 根据 f_{hs} 和 e_r 对降水重新赋值。

2.2 试验方案设计

方案 1 降水量不做任何变化,直接作为预报对象建立预报模型,并进行预报试验。

方案 2 当降水量为 0 时,若 $f_{hs} < 0$,则把降水值赋为 $f_{hs} \times e_r$;重新赋值后的时间序列作为预报对象建立预报模型,进行预报试验,预报结果中小于 0 的降水量样本重新赋值为 0。

方案 3 当降水量为 0 时,若 $f_{hs} < 0$,则把降水值重新赋为 $1.5 \times f_{hs} \times e_r$,重新赋值后的时间序列作为预报对象建立预报模型,进行预报试验,预报结果中小于 0 的降水量样本重新赋值为 0。

以上 3 个方案是定量降水预报,通常情况下人们把降水分级处理,并认为能很好地提高预报效果,为了比较定量降水预报与分级预报的预报效果,同时增加了试验方案 4。

方案 4 降水实况分为 5 级,分别代表无雨、小雨、中雨、大雨、暴雨及以上量级,具体处理过程请参阅文献[11]。

3 预报模型的建立

建立模型所用的资料为 2003 年和 2004 年 6—8 月的资料,预报因子为 12:00 起报的 T213 模式、德国模式、日本模式预报降水量和 T213 模式预报对流降水量共 16 个因子,对应的实况资料为 00:00 的 24 h 降水量。

4 个试验方案分别就同样的资料同样的方法建立

预报模型,对 2005 年 6—8 月北京、拉萨、乌鲁木齐、西安、上海、南昌和广州进行试报,并对试报结果进行检验,下面以 36 h 预报为例给出检验结果分析。

4 试验结果分析

下面从 TS 评分、空报率、漏报率以及预报偏差 4 个方面来对比分析不同降水处理方案预报效果的差异。

4.1 TS 评分

36 h 预报的小雨 TS 评分如图 1 所示,方案 1

的评分最低,乌鲁木齐、南昌和西安 TS 评分不到 0.3,说明直接对降水量进行预报,其预报效果是很差的。分级预报(方案 4)TS 评分有所提高,除乌鲁木齐外,都达到了 0.4 以上,但评分仍然低于方案 2 和方案 3。利用相对湿度对降水进行处理后的方案 2 和方案 3 评分最高,乌鲁木齐的评分分别达到了 0.68 和 0.70。方案 2 和方案 3 相比较方案 1 和方案 4 的预报效果都有明显的提高,尤其是乌鲁木齐从 0.26 和 0.38 提高到了 0.68 和 0.70,南昌从 0.29 和 0.43 提高到了 0.75 和 0.75。从整体情况看,方案 1 和方案 3 的评分差别不是很大。

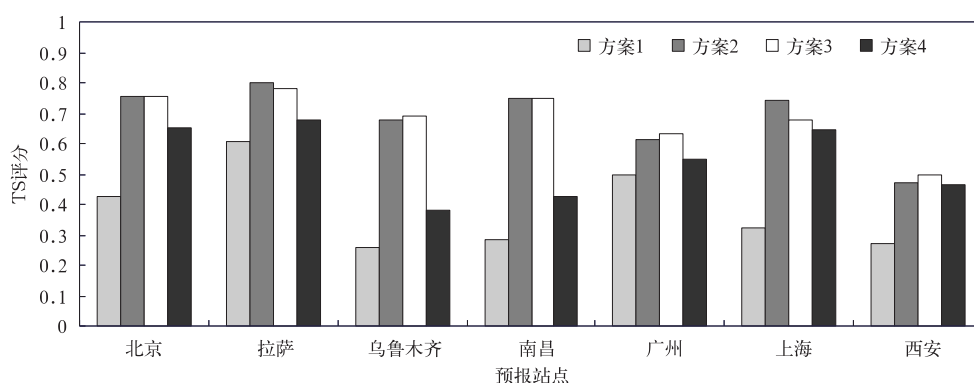


图 1 2005 年 6—8 月 36 h 降水预报小雨 TS 评分

Fig. 1 TS score of 36 h light rain forecast averaging from June to August in 2005

4.2 空报率

图 2 给出了 4 个处理方案的空报率:方案 1 的空报率最大,乌鲁木齐达到了 0.74,西安达到了 0.72,而南昌也达到了 0.71,空报率最小的拉萨也达到了 0.38,这表明方案 1 空报严重。方案 2 和方案 3 空

报率较小,除西安外,其余站点的空报率都在 0.3 以内,南昌方案 2 和方案 3 的空报率为 0。方案 4 的空报率从整体上看大于方案 2 和方案 3。说明相对于方案 4 而言,方案 2 和方案 3 对空报的情况也有所改进。

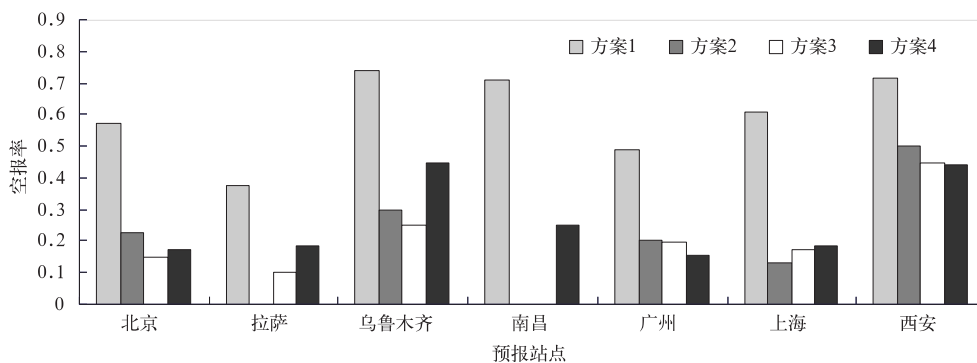


图 2 2005 年 6—8 月 36 h 降水预报小雨空报率

Fig. 2 No-hitting rate of 36 h light rain forecast averaging from June to August in 2005

4.3 漏报率

从图3中漏报率可以看到,方案1的漏报率最小,除西安的漏报率为0.15外,其余站点都小于0.1,北京和乌鲁木齐的漏报率为0,这和方案1空报率较大是一致的。由于方案1的空报率很大,即使它的漏报率很小,它的TS评分仍然很低。方案4的漏报率

最大,其中南昌达到了0.5,也就是说50%的降水没有报出来,而乌鲁木齐也达到了0.45;方案2和方案3的漏报率大于方案1而小于方案4,但方案2和方案3本身的漏报率并不是很大,都在0.3以内,由于方案2和方案3空报率和漏报率都小于方案4,所以它们的TS评分明显高于方案4。

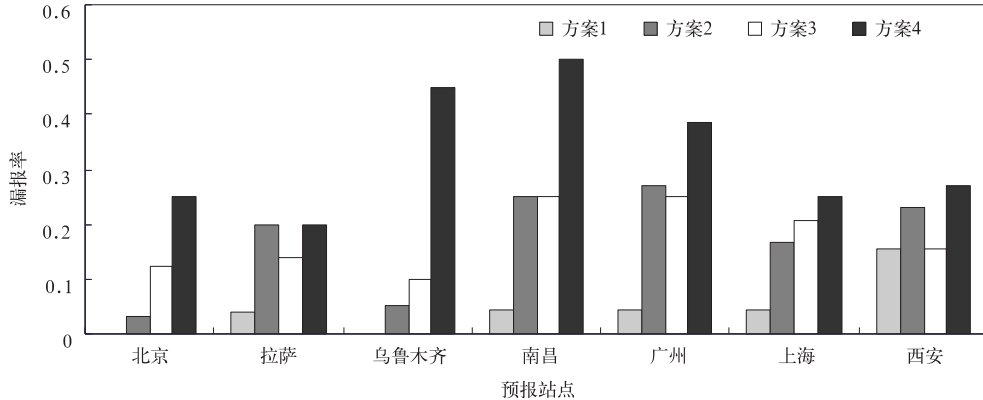


图3 2005年6—8月36h降水预报小雨漏报率

Fig. 3 Fault-hitting rate of 36 h light rain forecast averaging from June to August in 2005

4.4 预报偏差

预报偏差是衡量预报效果好坏的另一个重要指标,是预报样本数(区域)与实况发生该天气现象的样本数(区域)的比值,因此预报偏差接近或等于1最好,偏离1越大,说明预报样本数(区域)与实况样本数(区域)大小相差得越远。图4给出了上述7个站点小雨预报的预报偏差,从图中可以看到方案1

预报偏差较大,除拉萨为1.54、广州为1.86外,其余站点都大于2.0,其中预报偏差最大的乌鲁木齐达到了3.85,南昌达到了3.29,西安达到了3.0,说明这些站点预报样本数远远大于实况的样本数;方案2、方案3和方案4的预报偏差相近,整体上看没有明显差异,都接近1,说明预报样本数和实况样本数大小接近。

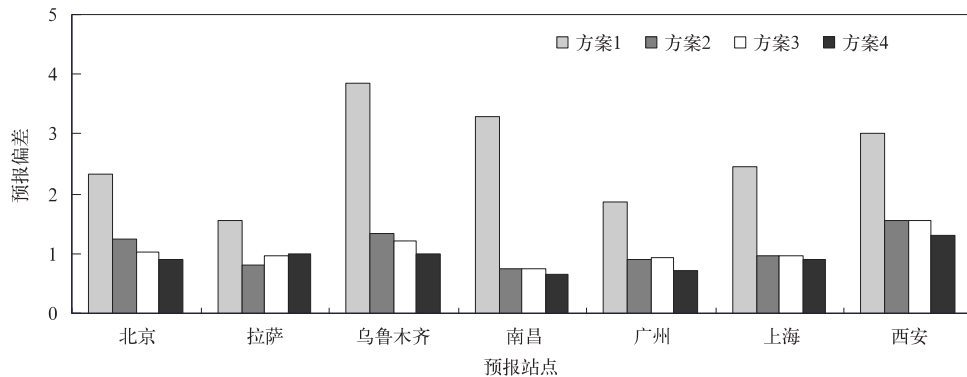


图4 2005年6—8月36h降水预报小雨预报偏差

Fig. 4 Forecast bias of 36 h light rain forecast averaging from June to August in 2005

以上的分析结果表明,利用相对湿度对零降水量的样本进行处理后,预报效果不论是从TS评分、空报率、漏报率还是从预报偏差上看都有提高,预报效果不但好于未对降水进行处理的预报方案,而且

也好于对降水进行分级处理的方案。

4.5 与模式预报评分的比较

前面的分析表明利用相对湿度对降水进行预处理后可以提高降水的预报效果。但是该处理方案与

T213 模式、德国模式和日本模式直接降水预报的预报效果相比又如何呢? 如果预报效果没有模式的直接预报效果好,那么所有的改进也毫无意义,因此,对 T213 模式、德国模式和日本模式的相应降水预报进行了插值,并对 2005 年 6 到 8 月的预报结果进行了检验,以方案 2 为例,给出了与模式的预报检验结果的对比分析。

从 TS 评分结果(图 5a)可以看到,所有站点方案 2 的评分明显高于 3 个模式直接降水预报的 TS 评分,除西安和广州外差别不是很大外,其他站点方案而较 3 个模式都有明显的优势。而 3 个模式评分

没有明显差别,T213 模式评分略大于德国模式和日本模式。这表明方案 2 的预报效果较单个模式直接输出预报有明显提高。

方案 2 的 TS 评分明显高于模式直接输出预报,那么空报和漏报的情况又如何呢? 由图 5b 可以看到 3 个模式空报率基本相当,日本模式略大于其他两个模式,而方案 2 的空报率较 3 个模式预报有明显降低,说明方案 2 的空报情况远远小于模式的直接预报。而漏报率(图略)总体上方案 2 较模式预报有所增加,但方案 2 的漏报率本身并不是很高,漏报率最大的广州站是 0.27,其次是西安,漏报率是 0.23,

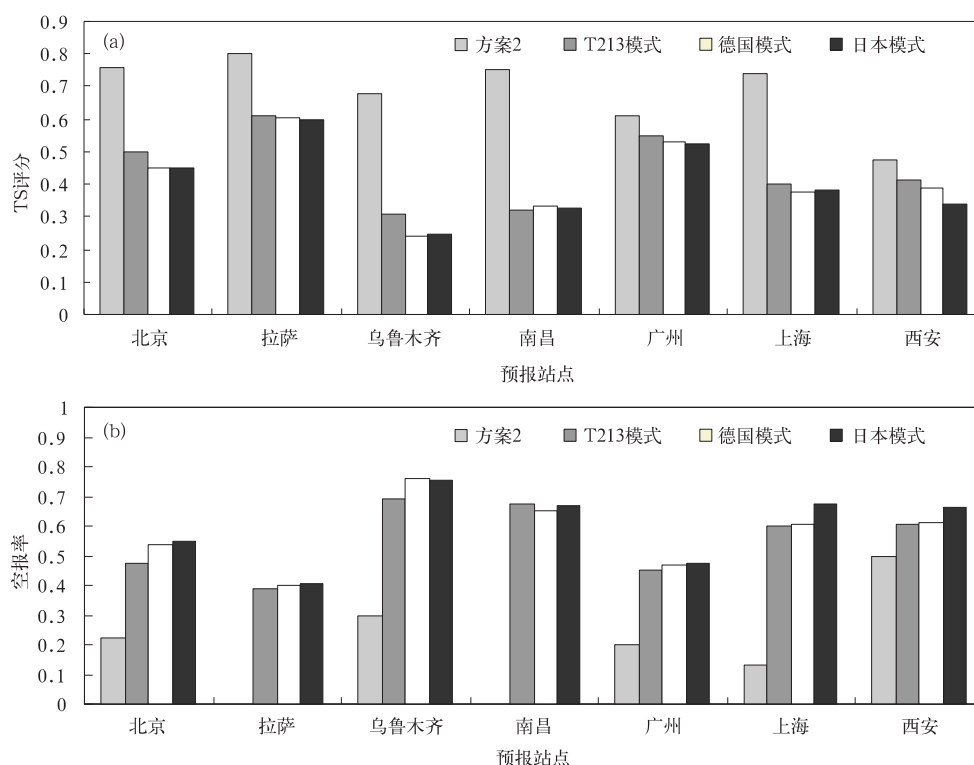


图 5 2005 年 6—8 月 36 h 降水预报小雨 TS 评分(a)和空报率(b)
Fig. 5 TS score(a)and fault-hitting rate (b) of 36 h light rain forecast
averaging from June to August in 2005

其他站点小于 0.2,而北京和乌鲁木齐漏报率小于 0.05,其中北京站的漏报率还小于 T213 模式和德国模式预报的漏报率。

从预报偏差(图略)也可以看到,方案 2 的预报偏差接近 1,而模式预报的预报偏差接近或大于 2,最大的达到 4.2,说明模式预报区域远大于实况区域。

与模式直接预报结果的对比分析表明,利用相对湿度对降水量进行适当的预处理,其预报效果同样明显好于模式直接预报。说明了该预处理方法切实可行,并且比较有效。

4.6 中雨及以上量级的检验结果

前面的检验结果是针对小雨预报的。图 6 则给出了方案 1、方案 2、T213 模式、德国模式和日本模式中雨、大雨的 TS 评分。从中雨的 TS 评分(图 6a)可以看到:7 个站中有 5 个站,方案 1 和方案 2 的评分较 3 个模式有非常明显的提高,而其他两个站的评分与 3 个模式相当,因此从整体上看经过释用以后中雨的预报效果有较大改进。两个释用方案的评分基本相当。大雨的评分情况与中雨相似。

中雨和大雨空报率(图略)情况一致,3 个模式

的空报率较高,而4个释用方案的空报率除广州和西安较低外,广州和西安4个方案的空报率与3个模式相当。漏报率(图略)的情况与空报率一样。这与TS评分是一致的。

分析表明对于中雨和大雨,从整体情况看,释用以后的效果明显好于模式预报。而利用湿度对降水量进行预处理的预报效果略好于不对降水量进行预处理和分级处理的预报效果。

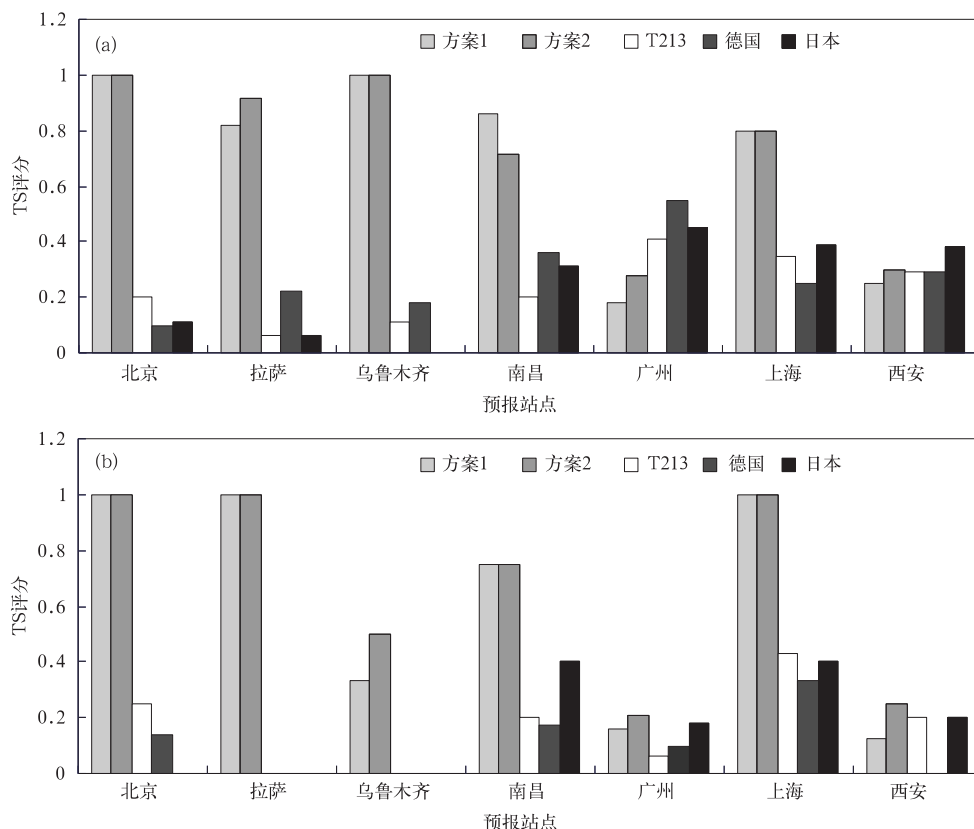


图6 2005年6—8月36h降水预报中雨(a)和大雨(b)TS评分

Fig.6 TS score of 36 h moderate rain (a) and heavy rain (b) forecast averaging from June to August in 2005

5 结论和讨论

以上分析表明:

1) 由于降水本身的特性,在建立预报模型之前,对降水量进行适当的处理,能够对预报结果产生较大的影响。

2) 利用相对湿度对零降水量的情况进行一定的预处理,再建立预报模型,预报效果有明显的提高。从TS评分、漏报率和空报率以及预报偏差综合来看,预报效果高于未对降水处理的预报效果,同时也好于对降水进行分级处理的预报效果;中雨和大雨的评分表明,方案2的预报效果略好于不对降水进行处理的预报效果,并好于对降水进行分级处理的

预报效果。

3) 与T213、德国和日本模式直接降水预报效果的比较表明,方案2的预报效果明显好于模式直接输出预报。

4) 从整体情况看,该降水处理方案对客观降水预报,尤其是定量的客观降水预报效果有改进,并且该处理方案简单可行。

本文对降水的预处理只是提供了一个思路和一种处理方法,不同站点还可根据具体情况采用不同的要素来对降水进行处理,例如云量等。

参考文献

- [1] 刘还珠,赵声蓉,陆志善,等. 国家气象中心气象要素的客观预报——MOS系统. 应用气象学报,2004,15(2):181-191.

- [2] 胡江林,涂松柏,冯光柳. 基于人工神经网络的暴雨预报方法探讨. 热带气象学报, 2003, 19(4): 423-428.
- [3] 刘爱鸣,潘宁,邹燕,等. 福建前汛期区域暴雨客观预报模型研究. 应用气象学报, 2003, 14(4): 420-429.
- [4] 赵声蓉. 多模式温度集成预报. 应用气象学报, 2005, 17(1): 52-58.
- [5] 张万诚,解明恩. 奇异值分解方法对降水的预测试验. 高原气象, 2002, 21(1): 103-107.
- [6] 宋海鸥,王永红,顾善齐,等. 应用 K 指数和 TOT 指数制作江苏中期降水预报的试验. 气象科学, 2002, 22(2): 242-246.
- [7] 陈力强,韩秀君,张立群. 基于 MM5 模式的站点降水预报释用方法研究. 气象科技, 2003, 31(5): 268-272.
- [8] 龚佃利,王以琳,谢考宪. 山东飞机增雨降水区分级预报方法研究. 应用气象学报, 2001, 12(增刊): 139-145.
- [9] 曹晓岗. 利用 T106 数值预报产品作江西暴雨动态落区预报. 江西气象科技, 1998, 21(1): 2-5.
- [10] 赵振宇,徐用懋. 模糊理论和神经网络的基础与应用. 北京: 清华大学出版社; 南宁: 广西科技出版社, 1996.
- [11] 赵声蓉. 多模式降水集成预报 // 章国材, 谢璞. 奥运气象保障技术研究. 北京: 气象出版社, 2004.

Preprocessing Precipitation in Objective Quantitative Forecast

Zhao Shengrong¹⁾ Pei Haiying²⁾

¹⁾ (National Meteorological Center, Beijing 100081)

²⁾ (Jiangsu Provincial Meteorological Observatory, Nanjing 210008)

Abstract

Quantitative precipitation forecast is much difficult in objective forecast. It is because that the precipitation is not a continuous variable, either not a normal distribution variable. Except that the precipitation has an obviously different characteristic to other elements that zero precipitation include many different situations. For example, black clouds blot out the sky but precipitation amount is zero. There are not many clouds on the sky and precipitation amount is also zero. These two situations are different, but the difference is not exhibited in the amount of precipitation. So the effectiveness of the forecast model is affected.

Preprocessing has important effects on forecast results. For the purpose of improving quantitative precipitation forecast effectiveness, a reasonable preprocessing scheme needs to be developed. Precipitation observation is preprocessed using relative humidity before modeling. The amount of precipitation is changed to different negative values when relative humidity is smaller than usual and precipitation amount equals zero. Based on the summer data of 2003 and 2004, forecast model of preprocessing precipitation and direct precipitation are developed respectively.

BP network method is used in the study, which is a kind of artificial neural network. The BP network is a back propagation network. It contains input layer, implicit layer and output layer. There can be one or more implicit layers on the BP network. Joint on the implicit layer is named implicit node. Input signals propagate to implicit nodes. Then signals of implicit nodes propagate to the next layer after the disposal of weights and operating functions. At last, the value on the output nodes is gotten. BP networks can be considered as a nonlinear projection from input layer to output layer. Sigmoid function is always taken as the operating function. Network weights of different nodes are obtained by training. The BP algorithm is used in training process.

Predictors are corresponding to precipitation amount forecast of operational global model of China National Meteorological Center, operational models of German Meteorological Administration and Japan Meteorological Agency. In order to avoid errors caused by interpolating, to every model, precipitation forecast on four grids around the stations are used as predictors instead of interpolated model precipitation forecasts at stations.

Test of different forecast results during 2005 summer of Beijing and other 5 stations indicate that precipitation forecasts are improved after preprocessing. Especially absent forecasts are reduced. So it can be concluded that this reprocessing method is simple and effective. But it only provides some ideas. In real settings, different element also can be used to reprocess precipitation amount according to different areas or different seasons in a year.

Key words: preprocessing for precipitation; absent forecast quotiety; leak forecast quotiety; forecast departure

中国气象科学研究院隆重举行成立 50 周年庆祝大会

2006 年 12 月 30 日上午,中国气象科学研究院隆重举行建院 50 周年庆祝大会。

中国气象局局长秦大河、中国科学院副院长李家洋应邀出席并发表热情洋溢的讲话,科技部基础司司长张先恩、国家自然科学基金委地学部主任陆则慰先后代表科技部部长徐冠华和国家自然科学基金委主任陈宜瑜宣读贺信。

中国气象局局长秦大河院士在讲话中肯定了气科院 50 年来坚持以国家需求为导向,积极开展大气探测、人工影响天气、灾害天气、天气与气候系统、生态环境与农业气象、数值模式、大气成分、大气雷电等多方面的研究,为提升我国气象业务水平,为国民经济建设做出的重要贡献。他说,经过 50 年的建设和发展,气科院已成为我国大气科学领域学科最多、规模最大的综合研究机构,气科院面向世界科技前沿,积极开展数值预报技术研究,为我国数值天气预报的发展奠定了基础。特别是研究建立了中国自己的新一代数值预报系统 GRAPES,为提高我国的数值天气预报准确率、缩小与发达国家的差距作出了一定的贡献;开展了东亚季风、暴雨、台风、热带气象、高原气象等方面的研究和试验,取得了一系列重要研究成果,丰富和发展了我国天气、特别是灾害天气的预报理论和方法;自主研发的 CAWS600 系列自动气象站已在全国气象业务台站配备;研制的城市污染预报系统已应用到中国大陆各省(区、市);新研制的区域污染系统正在全国推广。研究生教育事业得到长足的发展,培养出了一大批高素质的科研、业务和管理人才。面对气象事业快速发展的新形势,秦大河局长要求气科院在加强基础研究和应用基础研究的同时,要面向多轨道业务建设提供持续的、强有力的科技支撑,要加强国家重点实验室建设、加强科研与业务相结合,加强科技创新能力建设,引领中国气象局 8 个专业研究所的发展,努力打造科技创新团队。

中国科学院副院长李家洋院士对气科院建院五十周年表示了热烈祝贺,希望今后要进一步加强跨学科、跨领域、跨部门的合作,继续为中国气象事业的发展作贡献。

中国气象科学研究院院长张人禾在介绍气科院 50 年来改革和发展历程的报告中指出,经过 50 年的奋斗和磨砺,气科院已发展成为我国大气科学研究,特别是应用基础研究和技术开发的中坚和高层次人才的培养基地,并强调了中国气象局业务技术体制改革的推进,对气科院的工作赋予了新的内涵,为气科院今后的发展提供了难得的机遇和新的挑战。气科院将在着力提高科技创新能力,加强科研与业务的有机结合,拓展气象业务服务领域,从科技方面推动中国尽快从气象大国转变为气象强国。

中国气象局副局长许小峰主持了气科院建院 50 周年的庆祝大会,他希望气科院广大职工要认真学习、贯彻落实秦大河局长的讲话精神,在院领导班子的带领下,团结奋斗、勇于创新、扎实推进改革,不断促进中国气象科学研究院事业再上新台阶。

应邀参加庆祝活动的还有中国科学院所属研究所、中国农业科学院、中国地质科学研究院、中国林业科学研究院、国家海洋局所属研究所、有关大学以及中国气象局各直属单位、全国各省(区、市)气象局领导以及气科院职工、离退休职工代表、在读研究生代表等共 450 多人。期间,还收到了来自有关科研院所、院校以及中国气象局所属单位的贺电和贺信。

(中国气象科学研究院 黄幸媛)