

利用 ECMWF 产品对庆阳极端气温释用效果分析

吴爱敏^{1,2}, 路亚奇², 李祥科², 张洪芬², 徐 燕²

(1. 中国气象局兰州干旱气象研究所, 甘肃省干旱气候变化与减灾重点实验室,
甘肃 兰州 730020; 2. 甘肃庆阳市气象局, 甘肃 庆阳 745000)

摘 要: 利用 CMSVM 的回归方法, 用 2003~2007 年的欧洲数值预报格点资料和庆阳市 8 个自动站极端气温资料建立最高最低气温预报模型。2008 年业务运行效果检验评估表明: 5 d 最高、最低气温综合平均预报准确率达到 64% 和 71%, 对实时业务有较好的指导作用。最高、最低气温准确率随着预报时效的延长效果降低, 最低比最高气温的预报准确率高。最低气温预报效果春夏季好于秋冬季, 最高气温相反。最高、最低气温绝对误差和平均误差都随预报时效的延长增大, 最高比最低气温平均绝对误差大, 二者的平均误差接近, 均为正值, 预报值具有偏高的倾向。对 9 月明显变化过程的评估表明, 最高最低气温预报与实况演变的趋势相似。

关键词: 欧洲预报; 极端气温; 释用; 效果分析

中图分类号: P457.3

文献标识码: A

引 言

随着社会的快速发展, 对天气预报准确率的要求也越来越高, 精细化、较长时效的预报是我们解决的主要问题。近年新技术、新方法、新设备不断投入应用, 天气预报技术不断提高, 中国气象局相继开发了多种数值预报模式, 下发的信息不断增加, 市级台站加强了对数值预报的产品释用力度。如何从多种大气探测资料、数值预报模式等这些海量信息中获取可用于预报的关键信息, 通过机器学习就可能有效地解决这一问题。当我们面对数据而又缺乏理论模型时, 统计分析是最先采用的方法。多元回归、卡尔曼滤波方法等是建立在线性相关基础上的统计方法, 把它们用于具有非线性特征气象要素的预报上有局限性, 人工神经元模型通过给定数目的观测所得的解往往是局部最优, 在气象上的应用也存在不足。近年来发展起来的一种机器学习方法——支持向量机(Support Vector Machines 简称 SVM)方法为我们解决这一类问题提供了比较有效的手段。

对于气温预报, 国内外预报服务中根据不同的需求制作了各种精细化的温度预报产品, 开展了日极端气温和定时温度的预报, 数值预报产品中也直

接输出 2 m 高温预报, 由于模式本身误差, 初始边界条件等种种原因, 模式预报产品和实际气温状况存在较大差异, 各地针对数值预报模式开展气温的释用, 效果不尽一致。赵声蓉^[1]利用神经网络方法中的 BP 网络建立了中国 600 多个站的温度集成预报系统, 集成的温度预报结果明显优于 3 个模式单独的预报; 佟华^[2]总结了 T213L31 全球中期数值天气预报系统中 2 m 温度预报存在系统性偏低的问题; 陈豫英^[3]采用 2 种 MOS 统计方法, 预报宁夏 25 个测站 48 h 逐时温度, 比原 MM5 模式直接输出结果有明显提高; 马清^[4]利用类似 KALMAN 滤波的自适应误差订正法对中国国家气象中心、日本气象厅、美国国家环境预报中心、加拿大气象中心和澳大利亚—法国气象局的区域集合预报模式 2 m 温度预报做订正, 订正后温度预报的各项检验指标都显示出不同程度的改善; 马学款^[5]采用动态学习 BP 算法的人工神经网络建模, 在 ECWFM、T213 等模式数值预报产品释用基础上, 进行 1~7 d 逐日最高、最低气温模拟预测; 徐琳娜^[6]以 ECMWF 数值预报输出产品为基础, 对卡尔曼滤波(Klm)、最优子集回归(OSR)和岭回归(RR)3 种方法针对温度预报进行了探讨; 孙兰东^[7]运用 T106 模式数值产品历史资

收稿日期: 2009-03-04; 改回日期: 2009-07-23

基金项目: 甘肃省气象局多模式气温释用及预报集成项目(2009-07)资助

作者简介: 吴爱敏(1967-), 女, 河南温县人, 高级工程师, 主要从事中短期天气预报及服务. E-mail: qxjwam@sohu.com

料,用逐步回归方法,制作了甘肃省 19 个站春、夏、秋、冬 4 季的最高、最低温度的 120 h 内 5 个时效的预报方程;吴爱敏^[8-9]利用 HLAFS、T106 资料开展了庆阳气温降水预报,效果较好。以上研究利用不同模式、不同的统计方法对不同地区开展气温的预报,不同方法之间进行对比分析并尝试做集成预报。

ECMWF 数值预报模式性能稳定,积累资料多,评估检验表明 850 hPa 温度预报误差较小^[8],而且 ECMWF 850 hPa 温度格点资料能较好反映低层气温的变化。庆阳市海拔高度为 980 ~ 1 443 m,大部分地方超过了 1 200 m,850 hPa 接近地面,利用 CMSVM 统计方法对 ECMWF 850 hPa 格点温度进一步修正,可用来制作当地极端最高、最低气温的预报。因此选取 ECMWF 850 hPa 温度格点作为预报因子,利用中国气象局培训中心开发的支持向量机(CMSVM)统计预报软件,制作庆阳市逐月逐站极端气温预报模型,经不断优化,预报效果较好。与其他预报方法相比,该预报方法操作便捷,建模自动化、运行时间短、系统稳定、预报准确率较高,适合地市级气象台站数值预报产品释用工作的开展。

1 因子选用

1.1 资料

资料时段为 2003 ~ 2007 年,预报因子选取 ECMWF 数值预报模式格点,预报对象为庆阳市所属的 8 个自动站最高、最低气温,为了满足市级气象台每天下午制作预报的需要,选取每日 20 时 850 hPa 的温度格点资料,极端最高、最低气温资料为各自动站 2003 ~ 2007 年逐日的实况资料。

1.2 资料库的建立

预报因子选取 ECMWF 850 hPa 温度产品关键区 32.5 ~ 40°N,102.5 ~ 110°E 格点场上 16 个格点建立因子库,这样直接选取格点资料作为预报因子,一方面避免插值引起的误差和计算的复杂化,另一方面考虑了模式预报的空间特征。预报对象为最低、最高气温,根据各自出现的时间,选取靠近 ECMWF 资料的时次。即选取前一天 20 时 850 hPa 温度作为第 2 日最低温度的预报因子,选取预报当日的 20 时温度作为当日最高温度的预报因子。考虑到不同季节的影响,分别建立每月的预报模型。在制做每月的预报模型时,同时考虑季节的前后推移,将当月资料上下各跨 20 d,以确保当月温度预报模型的稳定性。例如建立 6 月资料库,预报因子和预

报对象的选取时段为 2003 ~ 2007 年每年的 5 月 11 日至 7 月 20 日。

资料库结构根据 CMSVM 统计预报软件要求设计(表 1),表 1 中第 1 列为预报对象,即极端最高或最低气温值,为了运算方便,建模时扩大了 10 倍,第 2 ~ 17 列为因子序号及 16 个格点因子值。

表 1 资料库结构

Tab. 1 The structure of information

最高气温 × 10	因子 1	因子 2	……	因子 16
234	1:18	2:17	……	16:21
220	1:18	2:16	……	16:21
214	1:18	2:16	……	16:23
231	1:19	2:21	……	16:19

2 预报模型的建立

2.1 支持向量机模型

SVM 方法的基本思路是:定义最优线性超平面,并把寻找最优线性超平面的算法归结为求解一个凸规划问题。进而基于 Mercer 核展开定理,通过非线性映射 φ ,把样本空间映射到一个高维乃至无穷维特征空间(Hilbert 空间),简单地说就是升维和线性化。升维,是把样本空间向高维空间做映射,在低维样本空间无法线性处理的样本集,在高维特征空间却可以通过一个线性超平面实现线性划分(或回归)。SVM 方法的核心是支持向量。SVM 建模方法,其本质就是通过对各种典型空间(支持向量)的充分表述来描述因子群与预报对象之间的关系,是一种基于事实的转导式推理。

2.2 确立核函数

以径向基函数(满足 Mercer 定理条件,又称高斯核,简记 RBF)作为核函数建立推理试验模型。径向基函数为: $K(x, x_i) = \exp(-r \|x - x_i\|^2)$,其中 x_i 作为支持向量机的样本因子向量; r 为核参数,求和运算只对支持向量机进行。

2.3 建模

利用建成的逐月资料库,建立逐月逐站点的最高、最低气温预报模型。按照 SVM 建模步骤,将资料库分为 3 部分:训练集、实验集、检验集,其中训练集占资料总数的 75%,实验集为 20%,检验集为 5%。采用中国气象局培训中心开发的 CMSVM 应

用软件建立回归模型。经过反复实验,把 T_s 评分最高的参数对应的支持向量机构造的推理模型作为最终确定的推理模型,把该模型参数推广应用,最终选定回归参数,进行建模,共建立逐月逐站点的最高、最低气温预报模型 $2 \times 12 \times 8$ 个。以庆阳市 4 月最低温度为例,所建立模型为:

```

2 # 核函数类型 -t
最优模型中核函数参数 -C 200
-1 # 最优模型中核函数参数 -d
0.006 # 最优模型中核函数参数 -g
1 # 最优模型中核函数参数 -s
1 # 最优模型中核函数参数 -r
-1 # 最优模型中核函数参数 -u
16 # 训练样本的特征空间的最高维数
522 # 训练样本的个数
261 # 支持向量的个数
-151.47416 # threshold b, 以下每行代表一个
支持向量(每行第一个实数代表 alpha * sign(y))
-44.9501 1:17 2:14 3:13 4:13 5:13 6:13 7:
12 8:12 9:12 10:14 11:12 12:12 13:16 14:14 15:14
16:13
.....

```

3 实时业务资料处理及业务运行

模型建立后,利用每天下发的 DVB—S 资料后台实时处理,定时自动读取前一日 20 时的欧洲数值预报不同时效的温度格点资料写入数据库,自动生成最高、最低气温预报模型中所需的数据文件。目前得到的欧洲数值预报时效为 168 h,在实际应用中为满足业务中对极端气温的精细化预报要求,选取 24 ~ 144 h 预报资料制作未来 5 d 最低温度预报,48 ~ 168 h 资料作为未来 5 d 最高气温预报。业务运行时,调入已建好的当月模型,对应实时处理好的欧洲预报因子数据,计算得出未来 5 d 逐县最高、最低气温预报结论,操作方便快捷,满足日常业务需求。

4 效果检验评估

对 2008 年的欧洲温度预报效果检验评估,评定 | 预报 - 实况 | ≤ 2 °C 的预报准确率 T_s , $T_s =$ 预报正确次数 / 预报总次数 $\times 100\%$, 5 d 120 h 综合评分按照甘肃省最新评分方法所要求的比例进行计算所得。

$$即 T_{s综合} = T_{s24} \times 40\% + T_{s48} \times 25\% + T_{s72} \times 20\%$$

$$+ T_{s96} \times 10\% + T_{s120} \times 5\%$$

4.1 庆阳市预报效果检验分析

2008 年 1 ~ 2 月庆阳气温比历年明显偏低,出现了近 30 a 不遇的异常低温天气,1 月平均气温比历年同期偏低 1 ~ 3 °C,是 1977 年以来最低,北部环县、华池是 1957 年有气象记录以来最低,创 1 月极端最低值。

CMSVM 建模所取的欧洲数值资料为 2003 ~ 2007 年,这 5 a 在全球气候变暖的大背景下,庆阳气温持续偏高,对 1 ~ 2 月出现的异常偏低气温模型概括率差,预报误差大,准确率较低,3 ~ 12 月相对稳定。表 2 为 2008 年 3 ~ 12 月预报模型对庆阳市各时次的逐月极端温度预报准确率评定,从评分结果可以看出,5 d 最高、最低气温综合平均预报准确率为 64% 和 71%,5 个时次中最低气温平均 > 64%,24 ~ 48 h 为 74%,效果较好。最高气温比最低气温差,24 ~ 96 h 内 4 个时次平均预报准确率 > 60%,120 h 为 51%。最高、最低气温准确率随着预报时效的延长效果降低。从以上分析可看出,最低气温比最高气温的预报准确率高,这与最高气温比最低气温变率大有很大关系。

表 2 2008 年 3 ~ 12 月庆阳市
最高、最低气温预报准确率(单位:%)
Tab.2 The forecast accuracy of the highest
and lowest temperature in Qingyang from
March to December, 2008 (Units:%)

项目	预报 时效	3 月	4 月	5 月	6 月	7 月	8 月	9 月	10 月	11 月	12 月	平均
最高 气温	24 h	40	67	76	55	73	61	57	83	83	75	67
	48 h	59	63	68	52	70	68	53	77	77	84	68
	72 h	54	59	48	45	70	68	50	70	70	65	60
	96 h	52	64	62	48	63	61	43	73	73	60	60
	120 h	58	56	28	45	63	52	47	63	63	40	51
	综合	50	64	65	51	70	64	53	77	77	72	64
最低 气温	24 h	83	93	83	79	77	61	60	70	70	65	74
	48 h	86	83	75	83	77	68	57	70	70	70	74
	72 h	68	76	59	79	73	68	67	70	70	60	69
	96 h	63	68	54	76	70	68	57	57	57	75	64
	120 h	65	81	52	59	63	61	47	57	57	65	67
	综合	78	84	72	79	75	65	60	68	68	66	71

从综合评定的结果看,最低气温预报准确率 3 ~ 7 月 > 72%,4 月预报效果最好,达 84%,其次是 6 月,为 79%,其中 4 月 24 h 的预报准确率为 93%,48

h 为 83% ,120 h 的预报为 81% ,这些时次的预报准确率甚至超过了预报员主观预报水平,对实时预报业务有很好的参考价值。9 月最低气温综合效果较差,仅为 60% ,由于 9 月下旬出现连阴雨天气,气温比历年同期偏低 1.7 °C ,中旬后期至下旬预报误差大导致全月整体预报效果下降。最高气温 10 ~12 月高于 72% ,10 月和 11 月检验效果最好,为 77% ,4 月份效果最差,仅为 50% ,4 月是冬季向夏季的过渡季节,此时冷暖空气交替,气温变率大,特别是最高气温跳跃大,所建预报模型准确率较差。从分析可看出,最低气温综合检验效果春夏季好于秋冬季,最高气温相反,秋冬季好于春夏季,这与日常预报员经验相近,春夏季升温快,最高气温变率大,秋冬季气温下降快,最低气温变率大,而且总体最高气温比最低气温变率大,预报模型的客观预报效果比较一致,变率大的季节预报效果差。

4.2 平均误差和绝对误差分析

预报的平均误差反映的是数值预报模式及释用后的系统误差,平均绝对误差是衡量预报效果的根本指标。用预报值减去实况值的差值,分析 2008 年 3 ~12 月预报模型制作的庆阳市最高、最低气温误差(表 3) ,5 d 平均最高、最低气温绝对误差分别为 2.01 和 1.65 ,平均误差均为 0.89。不论是最高还是最低气温,绝对误差和平均误差都随预报时效的延长增大,最高比最低气温平均绝对误差大,表明最低气温的预报效果整体好于最高气温,与表 2 中预报准确率的分析一致。但二者的平均误差却很接近,而且均为正值,说明所建模型的最高、最低气温系统误差为正,预报值高于天气实况值,具有预报偏高的倾向,业务应用中注意修正。

表 3 2008 年 3 ~12 月庆阳市最高、最低气温预报误差(单位:°C)

Tab.3 The prediction error of the highest and lowest temperature in Qingyang from March to December, 2008 (Units:°C)

项目	24 h	48 h	72 h	96 h	120 h	平均	
最高气温	平均绝对误差	1.76	1.76	2.0	2.05	2.48	2.01
	平均误差	0.75	0.83	0.87	0.92	1.09	0.89
最低气温	平均绝对误差	1.45	1.52	1.63	1.77	1.88	1.65
	平均误差	0.77	0.76	0.87	1.0	1.03	0.89

4.3 CMSVM 预报模型对温度明显变化过程预报能力分析

通过表 1 的分析,9 月最高、最低气温预报准确率较低,检验评估 9 月预报模型对明显温度变化过程的预报能力(图 1)。9 月有 3 次明显冷空气活动,分别是 2 ~3 日、7 ~9 日、24 ~27 日,这 3 次冷空气活动都伴有降水,最高、最低气温预报与实况演变的趋势相似,能反映冷空气活动,前半月最低气温误差较大,15 ~25 日最高气温误差大,均是预报比实况高。最高气温的最大误差出现在 20 日,最低气温的最大误差出现在 25 日。月初和上旬 2 次冷空气较弱,移动快,最低气温变化大,21 ~28 日出现了连阴雨过程,下旬平均气温比历年偏低 1.7 °C ,最高、最低气温的最大误差出现在连阴雨天气开始和后期,气温变化剧烈,特别是最高气温预报明显偏高。由于建模的历史资料很少出现气温偏低幅度较大的气候状况,概括率差,异常气候很难准确预报,造成预报误差大。其他月份对明显温度变化过程预报能力明显好于 9 月份,在此不详述。

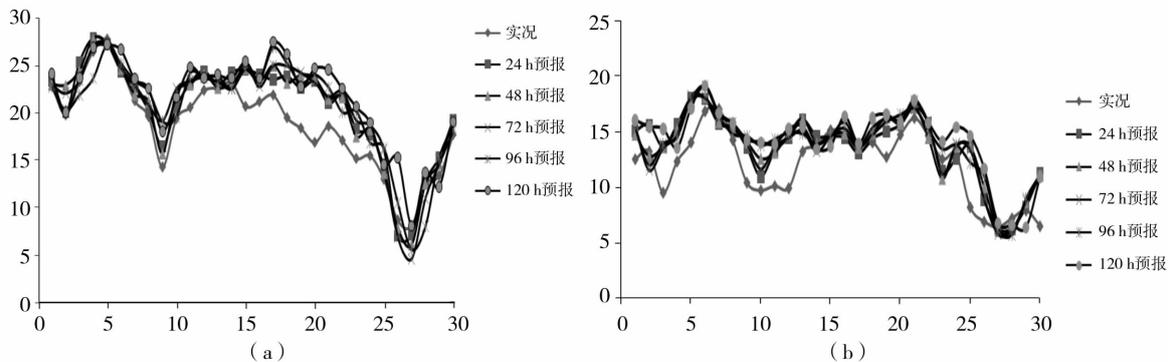


图 1 2008 年 9 月庆阳市最高气温(a)和最低气温(b)评估曲线图

Fig.1 The comparison between the prediction values by the CMSVM model and the observations of the highest (a) and lowest (b) temperature in September 2008 in Qingyang

4.4 庆阳市逐站点精细化预报能力分析

对 2008 年 3 ~ 12 月庆阳市所属的 7 县(环县、华池、庆城、镇原、合水、宁县、正宁)1 区(西峰)所有台站的极端气温预报进行效果评估,评估(表略)表明,120 h 精细化逐时效逐站点最高、最低气温预报准确率大部分时段超过 50%,部分时段在 60% 以上,有较好的指导意义。与前面分析的庆阳市检验评估效果相近,最低气温比最高气温整体准确率高,除个别月份个别站点外,西峰最高、最低气温的准确率均较其它站点高,和西峰距离近,地形相似的站点准确率高于远离的站点。分析原因,一是庆阳市南北跨度大,各县(区)的地理位置存在着较大的差异,西峰、合水和正宁处在陇东黄土高原塬区,地形相对平坦,华池、镇原、宁县、环县和庆城地处川道和山台地,特殊地形地貌造成了气温特殊性和变率大;二是 ECWFM 格点为 2.5×2.5 ,所选 $32.5 \sim 40^\circ\text{N}$, $102.5 \sim 110^\circ\text{E}$ 范围虽包括了庆阳市所有县(区),但以庆阳市所在地为着眼点进行格点选取,对于北部和南部有了偏差,这也是其他县比庆阳市预报效果差的原因之一。

5 方法的改进

ECMWF 产品 850 hPa 温度反映对流层低层的温度变化,模式输出的产品中已综合考虑了平流作用、垂直运动等因素的影响,通过和当地最高、最低气温建模,消除了系统误差,订正了地形等影响,预报指导意义好。但历史样本仅为 5 a,气候概括率较差,出现异常气候时段预报误差较大。建模时选取的因子要素单一,因为气温的变化和云量、云状、风等要素紧密相关,随着建模样本不断增加,进一步修正模型参数,完善优化预报因子,特别是增加如湿度、风等对温度影响显著的因子,改进模型稳定提高预报效果。

6 小结

用 ECMWF 850 hPa 温度单要素因子通过 SVM

建模制作 5 d 的精细化极端气温预报,和其他数值预报释用相比,方法简单,通过检验评估,对气温升降出现时间和变化幅度预报有着较好的指导意义。ECMWF 性能稳定,系统误差小,每日预报中,值班预报员根据该方法计算的温度值及其变化趋势,结合对本地天气气候变化的判断,极大地提高了最高、最低气温预报准确率。

2008 年 3 ~ 12 月的模型预报评分结果表明,5 d 最高、最低气温综合平均预报准确率达到 64% 和 71%,最高、最低气温准确率随着预报时效的延长效果降低,最低比最高气温的预报准确率高。最低气温预报效果春夏季好于秋冬季,最高气温相反。最高、最低气温绝对误差和平均误差都随预报时效的延长增大,最高气温比最低气温平均绝对误差大,二者的平均误差接近,均为正值,预报值具有偏高的倾向。对 9 月检验评估表明,最高、最低气温预报与实况演变趋势相似。庆阳市所属的其他台站的极端气温预报效果相近,但特殊地形地貌影响预报效果。

参考文献:

- [1] 赵声蓉. 多模式模式温度集成预报[J]. 应用气象学报, 2006, 17(01): 52 - 58.
- [2] 佟华, 姚明明, 王雨, 等. T213L31 全球中期数值天气预报系统 2 m 温度预报误差源分析[J]. 气象, 2006, 32(02): 52 - 57.
- [3] 陈豫英, 陈晓光, 马金仁, 等. 基于 MM5 模式的精细化 MOS 温度预报[J]. 干旱气象, 2005, 23(04): 52 - 56.
- [4] 马清, 龚建东, 李莉, 等. 超级集合预报的误差订正与集成研究[J]. 气象, 2008, 34(03): 42 - 48.
- [5] 马学款, 普布次仁, 唐叔乙, 等. 人工神经网络在西藏中短期温度预报中的应用[J]. 高原气象, 2007, 26(03): 491 - 495.
- [6] 徐琳娜, 冯汉中. 基于数值预报产品的温度释用方法比较[J]. 四川气象, 2006, 26(02): 4 - 5.
- [7] 孙兰东, 叶燕华, 王平鲁. 用 T106 数值产品制作甘肃省极端温度逐日滚动预报[J]. 甘肃气象, 2001, 19(03): 12 - 15.
- [8] 吴爱敏, 郭江勇. HLAFS 资料在短期降水、气温 MOS 预报方法中的应用[J]. 干旱气象, 2006, 24(02): 21 - 24.
- [9] 吴爱敏. 降水和温度 MOS 预报方法探讨[J]. 甘肃气象, 2002, 20(03): 2 - 5.

Effect Analysis on Interpretation of Extreme Temperature by ECMWF Products in Qingyang

WU Aimin^{1,2}; LU Yaqi², LI Xiangke², ZHANG Hongfen², XU Yan²

(1. *Institute of Arid Meteorology, CMA, Key Laboratory of Climate Change and Reducing Disaster of Gansu Province, Lanzhou 730020, China*; 2. *Qingyang Meteorological Bureau of Gansu Province, Qingyang 745000, China*)

Abstract: The prediction model for the highest and lowest temperature is established by the CMSVM regression method based on ECMWF data and the extreme temperature observations from 8 automatic weather stations in Qingyang of Gansu Province from 2003 to 2007. The effect test in 2008 operation forecast indicated that the rate of average forecast accuracy is 64% and 71% for the highest and lowest temperature in 5 days, which has the good instruction for the real-time weather service. The forecast accuracy rate of the highest and lowest temperature decreased with the forecast time increase, and the accuracy rate for the lowest temperature is higher than that of the highest temperature. The forecast effect for the lowest temperature is better in spring and summer than that in autumn and winter, but it is opposite for the highest temperature. The absolute and the average forecast errors for the highest and lowest temperature are added with the forecast time increase, and the absolute forecast error for the highest temperature is bigger than that of the lowest but the average error is close and positive. The forecast effect test in September of 2008 indicated that the forecast values presented the similar trend to real-time observations for the highest and lowest temperature.

Key words: ECMWF Products; the extreme temperature; interpretation; effect analysis



(上接第 275 页)

Analysis on a Regional Heavy Rain Process in East of Gansu Province

LIU Limin¹, Deqingcuomu², MENG Lixia¹

(1. *Lanzhou Central Meteorological Observatory, Lanzhou 730020, China*;
2. *Naqu Meteorological Bureau of Xizang Autonomous Region, Naqu 852000, China*)

Abstract: A regional rainstorm occurred in the east of Gansu Province from July 20 to 22, 2007 was analyzed from the circulation evolution, the main influence system and the features of physical quantity fields. Results show that the low-level jet provided sufficient moisture for the rainstorm, and the air stream convergence at low level and divergence at high level resulted in strong ascent motion, which was valuable for occurrence and maintain of the rainstorm, and ensured the development and maintenance of the precipitation cloud. The distribution of specific humidity and the pseudo-equivalent potential temperature has an indicative significance in prediction of rainstorm zone.

Key words: regional heavy rain; low vortex; low level jet stream; physical quantity field