

基于日较差 K 均值聚类的逐时温度预报

陶倩¹, 赵振宇², 张青珍¹, 刘东旭¹

(1. 辽宁省阜新蒙古族自治县气象局, 辽宁阜新 123100; 2. 辽宁省阜新市气象局, 辽宁阜新 123000)

摘要 采用 K 均值聚类分析方法, 将阜新蒙古族自治县气象局观测站 2006~2012 年 1 月的气温日较差进行分类, 以不同的气温日较差分型代表不同的天气状况, 并统计了相应分型下的逐小时温度变化幅度, 在此基础上对未来逐时温度做出客观预报。实况证明该方法对于实现时间上的精细预报, 具有一定的参考价值。

关键词 K 均值聚类分析; 气温日较差; 逐时温度; 预报

中图分类号 S161.2 **文献标识码** A **文章编号** 0517-6611(2013)07-03027-02

The Hourly Temperature Prediction Based on Daily Temperature Range K-means Clustering

TAO Qian et al (Fuxin Meteorological Bureau of Mongolia Autonomous County, Fuxin, Liaoning 123100)

Abstract By using k-means clustering analysis method, the daily temperature range of Fuxin Mongolia Autonomous County meteorological bureau in January 2006-2012 were classified. Different daily temperature range represents different weather conditions, and statistics was conducted on the corresponding hourly temperature variation range. On the basis of this, an objective forecast was made on hourly temperature in the future. The reality proved that the method has a certain reference for realizing precise forecasting.

Key words K-means clustering analysis; Daily temperature range; Hourly temperature; Forecasting

温度是一种连续、规律变化的气象因子, 温度变化受纬度、季节、地形、地势、下垫面性质、大气环流等多种因素影响。随着社会的进步、人民生活水平与生活需求的提高, 常规的最高气温与最低气温的预报已远远不能满足人类需求。当前人们更关注的是每日逐小时的温度, 因此迫切需求预报出每日逐小时温度。温度日较差是一日内最高气温与最低气温之差, 反映每日温度变化的幅度, 它能综合反映环流和天气要素对温度变化的影响。不同的日较差形势代表不同的天气状况, 高压控制下的晴好天气, 气温日较差较大; 低压影响下的阴雨天气, 日较差较小。而对于同一站点而言, 影响气温日较差的主要因素则为大气环流和天气条件, 因此, 可以用日较差来综合反映难以量化的大气环流和天气要素对气温日变化的影响, 在已有观测资料的基础上做出逐时气温预报。因此不同的气温日较差分型代表不同的天气状况。笔者结合 K 均值聚类分析方法的优点^[1-2], 将气温日较差分型, 统计出不同的日较差类型下温度逐小时变化幅度, 在此基础上客观地预报出未来逐小时温度。

1 资料与方法

1.1 资料选取 所用的资料是阜新蒙古族自治县气象局 2006~2012 年 1 月份自动站观测的逐小时温度, 以及人工观测的逐日日照。

1.2 分析方法

1.2.1 气温预报值的计算。在考虑气温日变化的基础上, 按日较差的值对资料样本分组, 分别求出各组的逐时气温变化幅度的气候平均值, 由数值预报模式预报的日较差对应预报日各时刻的逐时气温气候变量, 在观测初值上累加相应的气温变量, 从而得到预报日的逐时气温值。样本逐小时温

度变化计算公式为 $\Delta T_{ij} = T_{ij} - T_{ij+1}$ ($i = 1, 2, 3, \dots, 217; j = 1, 2, 3, \dots, 23$), 式中, i 为样本总天数, j 为逐日小时数。相应各型日较差对应的逐小时温度变化平均值计算公式为 $\Delta t_{ij} = \frac{1}{M} \sum_{i=1}^M \Delta T_{ij}$ ($i = 1, 2, 3, \dots, M; j = 1, 2, 3, \dots, 23$), 其中 M 为每个日较差分型所对应的样本天数。

1.2.2 K 均值聚类分析^[1-2]。采用 K 均值聚类分析方法对逐日气温日较差进行分类。 K 均值聚类分析原理是: 首先, 选择 n 个数值型变量参与聚类分析, 最后要求的聚类数为 k 个; 其次, 由系统选择 k 个观测量作为聚类的种子; 第三, 按照距离这些类中心的欧式距离最小的原则把所有观测量分派到各类重心所在的类别中去; 第四, 这样每类中可能由若干个样品, 计算每个类中各个变量的均值, 以此作为第 2 次迭代的中心; 然后根据这个中心重复第三、第四步, 直到中心的迭代标准达到最后要求。欧式距离公式为 $d_{ij} = \sqrt{\sum_{k=1}^n (t_{ik} - t_{jk})^2}$ 。

$$d_{ij} = \sqrt{\sum_{k=1}^n (t_{ik} - t_{jk})^2}$$

2 结果与分析

2.1 气温日较差与气温日变化 由阜新县气象局观测站 2006~2012 年 1 月份不同日较差范围内逐时温度变化幅度曲线(图 1)可见, 在不同的气温日较差下, 逐时温度变化振幅及变化趋势极不相同。因此, 可以在不同的日较差分型下, 根据逐时温度变化幅度和趋势做出未来逐时温度预报。

2.2 气温预报值的计算

2.2.1 日较差的分级。根据影响该站的大气环流以及出现频繁的天气条件, 将 2006~2012 年 1 月份逐日气温日较差分为 5 类, 5 个类的最终聚类中心分别为 19.9、17.0、13.8、7.3、10.3。由 5 个类中心间的距离(表 1)可以看出, 类 1 与类 2 间的距离最近, 为 2.995, 类 1 与类 4 间的距离最远, 为 12.630, 2006~2012 年 1 月份日较差分型中一型、二型、三型、四型、五型出现的天数分别为 25、53、56、33、50 d。

作者简介 陶倩(1984-), 女, 甘肃兰州人, 助理工程师, 从事地面气象观测工作, E-mail: xxkytq@163.com。

收稿日期 2013-02-21

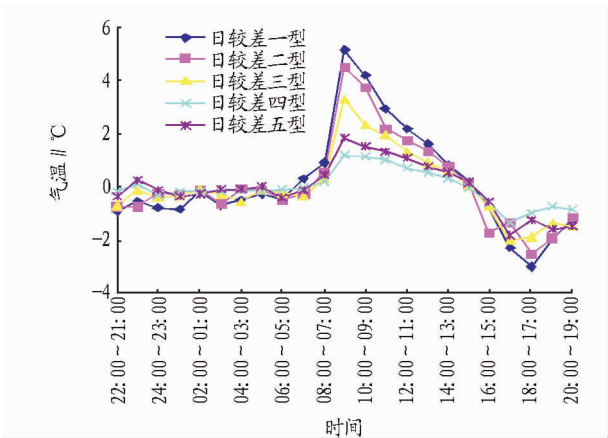


图1 2006~2012年1月不同气温日较差范围内逐时温度变化幅度

表1 最终聚类中心间的距离

聚类	1	2	3	4	5
1		2.995	6.141	12.630	9.610
2	2.995		3.146	9.635	6.615
3	6.141	3.146		6.489	3.469
4	12.630	9.635	6.489		3.020
5	9.610	6.615	3.469	3.020	

2.2.2 计算各级日较差对应的逐小时温度变化值。由2006~2012年1月份逐日逐时温度统计出各型日较差所对应的逐小时温度气候变化值(表2),根据数值预报做出的最高、最低温度值,可得到预报日所对应的日较差,将预报日的日较差与已经得到的5个日较差分型比较,可知道预报日的日较差分型,然后对照表2,得到相应日较差分型下的逐时温度变化值。

2.2.3 确定预报逐时气温的日界。地面观测规范以20:00~次日20:00为日界,逐时气温预报方法确定白天到夜间(08:00~次日08:00)、夜间到次日白天(20:00~次日20:00)2种时段的日界。当预报白天到夜间逐时气温时,选08:00气温观测值作为起报的初值因子;预报夜间到次日白天逐时气温时,选20:00气温观测值作为起报的初值因子。为了资料整理不繁琐,以21:00作为起报时刻来分析。

2.2.4 逐时气温预报值的计算。计算出逐时气温的1 h气候变量后,由数值预报做出的次日日最高气温与最低气温确定日较差。预报时刻的气温值用起报时刻气温观测值与逐时气温变化值的序列累加,就可依次计算出逐时气温值,其计算公式为 $T_k = T_0 + \sum_{j=1}^k \Delta t_{ij} (j=1,2,3,\cdots,23)$,式中, T_k 为逐

表2 5类气温日较差分型对应下的逐时温度变化值

时间	第1类	第2类	第3类	第4类	第5类
22:00~21:00	-0.95	-0.74	-0.71	-0.16	-0.37
23:00~22:00	-0.56	-0.75	-0.13	0.10	0.25
24:00~23:00	-0.78	-0.24	-0.43	-0.30	-0.13
01:00~00:00	-0.83	-0.33	-0.33	-0.16	-0.37
02:00~01:00	-0.18	-0.11	-0.08	-0.17	-0.27
03:00~02:00	-0.72	-0.66	-0.29	-0.20	-0.08
04:00~03:00	-0.53	-0.08	-0.57	-0.11	-0.07
05:00~04:00	-0.27	-0.14	-0.13	-0.18	0.01
06:00~05:00	-0.53	-0.50	-0.21	-0.11	-0.40
07:00~06:00	0.28	-0.29	-0.35	-0.06	-0.11
08:00~07:00	0.90	0.57	0.30	0.21	0.42
09:00~08:00	5.16	4.48	3.30	1.20	1.86
10:00~09:00	4.22	3.75	2.30	1.13	1.54
11:00~10:00	2.92	2.16	1.95	0.99	1.33
12:00~11:00	2.16	1.77	1.35	0.72	1.09
13:00~12:00	1.60	1.34	0.92	0.51	0.78
14:00~13:00	0.81	0.76	0.59	0.31	0.54
15:00~14:00	0.21	0.11	0.05	-0.05	0.14
16:00~15:00	-0.74	-1.74	-0.72	-0.62	-0.58
17:00~16:00	-2.28	-1.34	-2.02	-1.35	-1.82
18:00~17:00	-2.99	-2.52	-1.89	-0.99	-1.28
19:00~18:00	-1.92	-1.91	-1.37	-0.76	-1.60
20:00~19:00	-1.30	-1.19	-1.51	-0.85	-1.48

时气温预报值, T_0 为起报时刻气温观测值。

2.3 预报结果检验 为了检验预报方法的可用性,以2013年1月份21:00的观测值为初始温度,根据预报的日较差的分类,预报前1天22:00到第2日20:00的温度,检验结果发现预报温度与实际温度偏差在0℃左右的占18.50%,偏差<1℃左右的累计占36.7%,偏差<2℃左右的累计占60.8%。由此可见,该方法对预报逐时温度具有一定的参考价值。

3 小结与讨论

应用日较差的分级客观地预报逐时温度,当连续出现比较稳定的天气时,地面逐时气温预报准确率较高,而在一日内遇有突发性天气出现时气温预报误差较大,反映出统计方法的固有缺陷。随着样本的增加及引入其他的影响因子,将进一步提高其稳定性与准确率。

参考文献

[1] 黄嘉佑. 气象统计分析与预报方法[M]. 3版. 北京:气象出版社,2004.
[2] 张强. 统计方法讲座——(三)显著性差异[J]. 气象,1977(7):26~29.
[3] 宛敏渭. 中国自然历编[M]. 北京:科学出版社,1986~1987:1~125.
[4] 王植,刘世荣. 全球环境变化对植物物候的影响[J]. 沈阳农业大学学报:社会科学版,2007,9(3):350~353.
[5] 史彩霞,李耀先,李莉,等. WEB和GIS的“自然物候与气候变化关系的研究”系统平台的设计与开发[J]. 安徽农业科学,2012,40(1):1~4.
[6] 丁抗抗,高庆先,李辑. 我国植物物候变化及对气候变化的响应综述[J]. 安徽农业科学,2010,38(14):7414~7417.
[7] 祁如英,严进瑞,王启兰. 青海车前草物候变化及其对气候变化的响应[J]. 中国农业气象,2006,27(1):41~45.

(上接第3026页)

参考文献

[1] 陆佩玲,于强,贺庆棠. 植物物候对气候变化的响应[J]. 生态学报,2006,26(3):923~929.
[2] 蒋菊芳,王鹤龄,魏育国,等. 河西走廊东部不同类型植物物候对气候变化的响应[J]. 中国农业气象,2011,32(4):543~549.
[3] 李静,王秀珍. 浙江省自然物候空统计分析[J]. 杭州师范大学学报:自然科学版,2012,11(2):107~114.
[4] 陆置. 车前草研究述论[J]. 南阳师范学院学报,2011,10(6):58~62.