

卡尔曼滤波在短期气候预测中的应用^①

陈 雷 刘开福 李 英

(甘肃省武威地区气象台,武威 733000)

提 要

用 PRESS 准则和最优子集回归选取预测因子,建立河西走廊东部月平均气温的卡尔曼滤波预测模型,经实际使用发现,该预测方法的精度较高,稳定性较好,为短期气候预测提供了一种新的思路和方法。

关键词: 卡尔曼滤波 短期气候预测 月平均气温

引 言

卡尔曼滤波是一种统计估算方法,是数学家卡尔曼于 1960 年创立的。它通过利用前一时刻的预报误差反馈到原来的预报方程,并及时修正预报方程系数,以此提高下一刻的预报精度。它的预报对象一般为具有线性变化特征的连续性变量。目前该方法在短期天气预报和数值预报产品释用中,越来越被广泛地应用^[1~3],然而在短期气候预测(长期天气预报)中应用仍处于初级阶段。本文介绍用卡尔曼滤波方法建立河西走廊东部月平均气温的可变预测模型。

1 资料选取及处理

1.1 资料选取

选取了河西走廊东部武威、永昌、民勤、古浪、乌鞘岭 5 站(1961~1997 年)的月平均气温为预测对象;选取了 5 站(1960~1997 年)的地面月平均资料,民勤(1960~1997 年)的 500hPa 高空的月平均高度、温度、露点资料,74 个环流特征量的月平均资料为预测因子。

1.2 资料处理

将预测对象和预测因子按月进行极差标准化处理,各月的数值均在 $[0, 1]$ 范围,从而达到消除各月之间因数值相差悬殊而造成的

预测不稳定性。然后将经过极差标准化处理的月平均气温资料,按顺月顺年排列,作为预报对象量,并将其划分为三组样本,1961~1975 年 15 年 180 个月的资料作为依赖样本,1976~1990 年 15 年 180 个月的资料作为独立样本,1991~1997 年 7 年 84 个月为检验试报样本;把经过极差标准化处理过的武威等 5 站的资料及环流特征量也按顺月顺年排列,作为预报因子量。

1.3 预测因子的筛选

对一千多个预测因子,用 PRESS 准则^[4]初选出 132 个因子,再分次分步用逐步回归精选出 13 个因子,最后用最优子集回归^[5]确定出 8 个预报因子,它们分别为:

x_1 : 从 1960 年 7 月起按顺月顺年排列的北半球极涡强度指数,

x_2 : 从 1960 年 7 月起按顺月顺年排列的西藏高原(30~40°N, 75~105°E)环流指数,

x_3 : 从 1960 年 4 月起按顺月顺年排列的武威 10cm 月平均地温,

x_4 : 从 1960 年 5 月起按顺月顺年排列的民勤 500hPa 月平均高度,

x_5 : 从 1960 年 6 月起按顺月顺年排列

① 本文得到甘肃省气象局青年科技基金资助。

的民勤 500hPa 月平均高度,

x_6 : 从 1960 年 2 月起按顺月顺年排列的乌鞘岭月平均最高气温,

x_7 : 从 1960 年 7 月起按顺月顺年排列的武威月平均最低气温,

x_8 : 从 1960 年 12 月起按顺月顺年排列的武威月降水量,

预报对象为:

$y_1 \sim y_5$ 分别为从 1961 年 1 月起按顺月顺年排列的武威、永昌、民勤、古浪和包鞘

岭等 5 站的月平均气温。

2 卡尔曼滤波方法

卡尔曼滤波的预报对象一般为具有线性变化特征的连续性变量。通过计算分析 5 站顺月顺年的月平均气温序列的自相关系数 R 和预测平方和系数 R_p , 发现与落后 1~12 个月的 PRESS 值和相关系数大部分较好(表 1), 表明月平均气温的变化具有较好的连续性。

表 1 5 站月平均气温序列的自相关系数和预测平方和

落后月数	武威		永昌		民勤		古浪		乌鞘岭	
	R_p	R	R_p	R	R_p	R	R_p	R	R_p	R
1	0.97	0.99	0.96	0.98	0.97	0.99	0.96	0.98	0.95	0.98
2	0.72	0.85	0.72	0.85	0.72	0.85	0.72	0.85	0.71	0.84
3	0.23	0.49	0.23	0.49	0.23	0.49	0.23	0.49	0.23	0.49
5	0.24	-0.49	0.23	-0.49	0.23	-0.49	0.23	-0.49	0.23	-0.49
6	0.71	-0.84	0.70	-0.84	0.71	-0.84	0.70	-0.84	0.70	-0.84
7	0.94	-0.97	0.94	-0.97	0.94	-0.97	0.94	-0.97	0.94	-0.97
8	0.70	-0.84	0.70	-0.84	0.71	-0.84	0.70	-0.84	0.70	-0.84
9	0.23	-0.49	0.23	-0.49	0.23	-0.49	0.23	-0.49	0.23	-0.48
11	0.23	0.49	0.23	0.49	0.23	0.49	0.23	0.49	0.23	0.49
12	0.73	0.85	0.72	0.85	0.73	0.85	0.72	0.85	0.71	0.85

2.1 预测模型的建立

把 180 个极差标准化处理过的依赖样本作为预报对象, 用多元回归建立量测方程, 其复相关系 R_1 , 显著性检验 F , 残差平方和 Q 见表 2。

武威的量测方程:

$$\hat{y}_1 = 0.3686 + 0.0057x_1 + 0.1101x_2 + 0.1373x_3 + 0.219x_4 + 0.1109x_5 - 0.1486x_6 - 0.1193x_7 - 0.1005x_8$$

永昌的量测方程:

$$\hat{y}_2 = 0.3956 + 0.0331x_1 + 0.1082x_2 + 0.111x_3 + 0.1542x_4 + 0.1236x_5 - 0.1338x_6 - 0.0822x_7 - 0.0609x_8$$

民勤的量测方程:

$$\hat{y}_3 = 0.3092 + 0.0356x_1 + 0.1021x_2 + 0.1257x_3 + 0.2072x_4 + 0.1179x_5$$

$$- 0.0717x_6 - 0.1007x_7 - 0.0662x_8$$

古浪的量测方程:

$$\hat{y}_4 = 0.3968 + 0.0321x_1 + 0.1085x_2 + 0.0925x_3 + 0.1207x_4 + 0.1479x_5 - 0.0871x_6 - 0.1117x_7 - 0.0756x_8$$

乌鞘岭的量测方程:

$$\hat{y}_5 = 0.4535 + 0.0187x_1 + 0.1097x_2 + 0.0368x_3 + 0.1008x_4 + 0.1666x_5 + 0.1615x_6 - 0.1134x_7 + 0.1099x_8$$

表 2 5 站依赖样本的量测方程特征值

站名	复相关系数 R	显著性检验 F	残差平方和 Q
武威	0.401	4.099	9.296
永昌	0.305	2.194	11.169
民勤	0.351	2.995	9.876
古浪	0.348	2.845	9.997
乌鞘岭	0.334	2.689	10.567

把 180 个极差标准化处理过独立样本作

为预报对象,用多元回归建立量测方程,其复相关系数 R_2 ,显著性检验 F ,残差平方和 Q 见表3。

表3 5站独立样本的量测方程特征值

站名	复相关系数 R_2	显著性检验 F	残差平方和 Q
武威	0.426	4.728	8.988
永昌	0.357	3.121	9.806
民勤	0.380	3.614	8.620
古浪	0.363	3.285	9.657
乌鞘岭	0.368	3.340	9.599

武威的量测方程:

$$\hat{y}_1 = 0.5361 - 0.2857x_1 + 0.1516x_2 + 0.135x_3 + 0.0431x_4 + 0.0959x_5 - 0.0738x_6 - 0.1241x_7 - 0.1138x_8$$

永昌的量测方程:

$$\hat{y}_2 = 0.5808 - 0.267x_1 + 0.1327x_2 + 0.1025x_3 + 0.0374x_4 + 0.0804x_5 - 0.0542x_6 - 0.1167x_7 - 0.028x_8$$

民勤的量测方程:

$$\hat{y}_3 = 0.5443 - 2165x_1 + 0.0875x_2 + 0.1451x_3 + 0.0927x_4 + 0.0769x_5 - 0.0458x_6 - 0.0928x_7 - 0.1008x_8$$

古浪的量测方程:

$$\hat{y}_4 = 0.4836 - 0.2193x_1 + 0.1487x_2 + 0.1109x_3 + 0.0312x_4 + 0.0979x_5 - 0.0517x_6 - 0.1021x_7 - 0.0315x_8$$

乌鞘岭的量测方程:

$$\hat{y}_5 = 0.3802 - 0.175x_1 + 0.2184x_2 + 0.1427x_3 + 0.0271x_4 + 0.1251x_5 + 0.0019x_6 - 0.0788x_7 - 0.0947x_8$$

根据两组方程位相差 $\Delta T = 180$ 和残差 Q ,可算出动态噪声方差阵 W ,以及白噪声

V.卡尔曼滤波递推公式:

$$\hat{Y}_t = X_t \hat{\beta}_{t-1} \quad (1)$$

$$R_t = C_{t-1} + W \quad (2)$$

$$\sigma_t = X_t R_t X_t^T + V \quad (3)$$

$$A_t = R_t X_t^T \sigma_t^{-1} \quad (4)$$

$$\hat{\beta}_t = \hat{\beta}_{t-1} + A_t (Y_t - \hat{Y}_t) \quad (5)$$

$$C_t = R_t - A_t \sigma_t A_t^T \quad (6)$$

方程(1)是预报方程, \hat{Y}_t 为预报对象, X_t 为预报因子, $\hat{\beta}_{t-1}$ 为方程回归系数估算值。

方程(2)中的 R_t 为外推值 $\hat{\beta}_{t-1}$ 的误差方差阵, C_{t-1} 为滤波值 $\hat{\beta}_{t-1}$ 的误差方差阵、 W 为动态噪声方差阵。

方程(3)中的 σ_t 为预报误差方差阵, X_t^T 为预报因子 X_t 的转置矩阵, V 为白噪声。

方程(4)中的 A_t 为增益矩阵, σ_t^{-1} 为 σ_t 的逆矩阵。

方程(5)为系数 $\hat{\beta}_t$ 的订正方程, Y_t 为预报对象实测值。

方程(6)为计算 C_t 方程。

根据式(1)~(6)组成的递推系统,通过不断更新预报方程的系数建立起了月平均气温的预测模型。

3 效果检验

通过对独立样本的递推与试报样本的检验,22年264个样本,5站平均气温预报的总绝对误差、平均绝对误差、独立样本的平均绝对误差、试报样本的平均绝对误差、最大绝对误差等情况见表4。由于试报样本数较大,现只列出武威1998~1999年逐月平均气温实报情况(表5)。由此可见,卡尔曼滤波预测模型的预报精度较高,稳定性较好。

表4 5站月平均气温预报效果检验(单位:℃)

站名	总绝对误差	平均绝对误差	依赖样本平均绝对误差	试报样本平均绝对误差	最大绝对误差	总预报准确率	依赖样本准确率	试报样本准确率
武威	339.0	1.18	1.2	1.1	5.0	0.71	0.70	0.73
永昌	373.1	1.29	1.3	1.2	5.4	0.67	0.65	0.71
民勤	391.9	1.36	1.4	1.3	5.8	0.69	0.68	0.71
古浪	368.5	1.28	1.3	1.3	5.2	0.67	0.66	0.69
乌鞘岭	362.9	1.26	1.3	1.2	4.7	0.68	0.66	0.72

表5 1998~1999年武威站月平均气温实报效果检验(单位:℃)

月份	1	2	3	4	5	6	7	8	9	10	11	12	
1998年	预报值	-8.7	-5.0	4.7	10.9	18.0	18.6	23.4	21.3	15.2	10.0	2.6	-2.7
	实况值	-8.7	-1.1	2.7	13.4	15.2	21.2	22.6	21.0	16.9	9.1	3.1	-2.1
	误差	0.0	-3.9	2.0	-2.5	2.8	-2.6	0.8	0.3	-1.7	0.9	-0.5	-0.6
	平均值	-8.7	-5.0	2.6	10.0	15.8	19.9	21.9	20.5	14.9	7.8	-0.5	-6.9
	评定	√	×	×	√	×	×	√	√	√	√	√	√
1999年	预报值	-4.8	-3.4	4.1	11.5	18.0	19.6	23.4	21.8	15.6	10.0	2.5	-5.5
	实况值	-5.8	-2.3	3.7	12.6	16.3	21.4	22.7	21.7	17.5	9.5	1.0	-3.7
	误差	1.0	-1.1	0.4	-1.1	1.7	-1.8	0.7	0.1	-1.9	0.5	1.5	-1.8
	平均值	-8.7	-5.0	2.6	10.0	15.8	19.9	21.9	20.5	14.9	7.8	-0.5	-6.9
	评定	√	√	√	√	√	×	×	√	√	√	√	√

4 结束语

用 PRESS 准则粗选预报因子,最优子集精选因子,克服了其它方法筛选因子的不稳定性,从而使预测因子达到全局最优,得到的量测方程模型也较合理。

由于卡尔曼滤波在短期气候预测中的应用,尚属初级阶段。经试报和实报发现,本方法对月平均气温有较好的预测能力。但仍存在个别月份误差较大,这可能与模型的系列参数订正滞后有关,还需进一步分析改进。

致谢: 本文得到甘肃省气象局朱炳瑗研究员的指导和帮助,在此表示衷心感谢。

参考文献

- 1 黄嘉佑,谢庄. 卡尔曼滤波在天气预报中的应用. 气象, 1993,19(4):3~7.
- 2 陆如华,徐传玉,张玲等. 卡尔曼滤波在天气预报中的应用技术. 数值预报产品评价公报,1996,5~6 双月刊:28~36.
- 3 金荣花. 卡尔曼滤波技术在中期候平均温度预报中的应用. 数值预报产品评价公报,1996,9~10 双月刊:21~25.
- 4 俞善贤. 一个着眼于预测能力及稳定性的因子普查方法. 气象,1991,17(9):40~43.
- 5 甘肃省气象局科教处. 短期气候预测方法. 兰州,1997年7月:6~12.

Application of Kalman Filter Method to Short-range Climatic Prediction

Chen Lei Liu Kaifu Li Ying

(Wuwei Meteorological Observatory, Gansu Province 733000)

Abstract

Kalman filter method based on the PRESS criterion and optimal subset regression is applied to predict the monthly average temperature in the east of Hexi area. Results show that the prediction modeling of Kalman filter method is better and steadier. It provides a new method for the short-range climatic prediction.

Key Words: Kalman filter short-range climatic prediction monthly average temperature