

杨旭 袁国恩 (沈阳区域气象中心研究所)

在使用长序列气象历史资料中,我们往往会遇到解放前历史资料中出现缺测而间断的现象,感到使用不便。为了能利用宝贵的长序列资料,必须对缺测时段的资料进行弥补,并尽可能符合实际情况。以前人们都是采用线性的统计方法建立回归方程来达到这一目的,然而实际情况中相邻两站气象要素的相关性大多是非线性的,因此过去所采用的线性订正法偏差较大。经对沈阳、大连、丹东、营口和朝阳等站温度和降水资料的插补,结果证明非线性回归方程比线性回归方程不论在历史拟合方面还是统计检验效果上都有明显提高,因此插补的数值更切合实际。

一、因子模型

以往人们即使考虑非线性相关因子模型时也只是 x^{-1} 、 $x^{1/2}$ 、 x^2 这几种特殊情况,然而实际情况下 x^a 中 a 是在 $(-\infty < a < +\infty)$ 范围内取值的,这样考虑对相关因子之间的非线性关系才更具有普遍的意义,更接近于客观实际。我们使用非线性方

法将因子转换为以下3种模型。

1. 幂函数型: $Y = x^a$ ($-\infty < a < +\infty$)
当 $a = 1$ 时为线性函数。
2. 指数函数型: $y = e^{ax}$ ($-\infty < a < +\infty$)
3. 对数函数型: $y = \ln x$

因为第1、3两种类型要求 $x > 0$,所以首先必须将因子数据进行极差正规化处理:

$$X_i' = \frac{x_i - \min_{1 \leq i \leq N} \{x_i\}}{\max_{1 \leq i \leq N} \{x_i\} - \min_{1 \leq i \leq N} \{x_i\}}$$

使 $0 \leq x' \leq 1$ 。

然后,利用计算相关系数公式判断因子 y 与 x 相互关系的好坏,对于第3种类型可直接计算出相关系数 $|R_3|$,但对于第1、2种类型多有一个参数 a ,要计算出使两个因子 y 与 x 相关系数最大时的 a 值,这里采用单变量函数寻优的外推内插法,求得 a 值后便可计算出第1、2种类型的相关系数 $|R_1|$ 、 $|R_2|$ 。得到这3种类型的相关系数后,进行比较,从中选出 $|R_g|$ 最大的一个,则确定该因子为 g 类。

二、实用方程的建立

为了插补40年代沈阳各月气温和降水的缺测历史资料。限于所掌握的资料,以《抚顺气象三十年》的温度和降水资料作为因子,分别取沈阳的气温、降水作为预报量,缺测段为待报量,分别建立各月的温度、降水方程。为证实非线性回归方程优于线性回归方程,用非线性回归方程插补的资料更切合实际,又从中选出7月份温度和降水两个回归方程和方程检验结果进行线性和非线性的比较,见表1和表2。

从表1—2中可见,不但非线性相关系数比线性相关系数高,而且拟合平均绝对误差非线性回归方程比线性回归方程要小,其他各种检验效果都比线性方程要好。

表 1 7 月份沈阳温度线性和非线性回归方程检验结果

方程类型	回归方程	因子类型	R	Q	F	E
非线性	$\hat{Y} = 13.108 + 9.62872 \left[\frac{(x-8.3)}{18.3} \right]^{0.4979}$	$x' = x^a$	0.9541	2.1877	223.47	0.2458
线性	$\hat{Y} = 18.859 + 3.995 \left[\frac{(x-8.3)}{18.3} \right]$	$x' = x$	0.9529	2.2446	217.25	0.2580

表 2 7 月份沈阳降水线性和非线性回归方程检验结果

方程类型	回归方程	因子类型	R	Q	F	E
非线性	$\hat{Y} = 1414.13 - 1491.81 \times e^{-0.6 \left(\frac{x-1.3}{33.9} \right)}$	$x' = e^{ax}$	0.7090	306743.4	23.2446	87.315
线性	$\hat{Y} = 243.64 + 378.263 \left[\frac{x-1.3}{33.9} \right]$	$x' = x$	0.7058	309552.9	22.8249	88.3385

注: R 为相关系数, Q 为残差平方和, F 为检验值, E 为平均绝对误差。

表 3、表 4 中的数据就是用非线性回归方程插补的沈阳 40 年代缺测温度和降水资料。由于目前还没有发现有人公布这一段缺测资

料的插补数值, 因此将此组数据提供给读者参考使用, 作为本文的结束。

参 考 文 献 (略)

表 3 温度插补值

年代	温 度											
	1	2	3	4	5	6	7	8	9	10	11	12
1942					15.8	23.2	24.9	23.2	17.3	10.5	0.0	-9.7
1943	-14.5	-7.5	2.1	10.1	17.6	22.6	26.5	25.1	18.0	10.4	0.3	-7.0
1944	-12.1	-7.3	0.9	10.4	16.3	22.1	26.5	23.9	19.1	10.3	-1.1	-11.7
1945	-14.5	-12.4	-0.5	11.3	15.3	21.7	25.9	25.7	17.8	10.5	-0.4	-9.1
1946	-12.8	-9.2	0.2	10.0	16.7	22.1	25.4	24.2	17.3	10.1	-0.8	-9.4
1948							25.3	25.0	17.0	7.3	-0.8	-4.9

表 4 降水插补值

年代	降 水 (0.1 mm)											
	1	2	3	4	5	6	7	8	9	10	11	12
1943	22	70	96	356	822	365	845	1256	379	193	268	194
1944	5	7	271	91	993	2357	1249	2106	574	479	277	11
1945	63	49	230	320	833	826	1852	2455	551	333	222	93
1946	69	72	207	348	750	788	1912	1759	551	333	222	93
1948							1106	1107	475	289	69	41