

# 应用系统建模法对南海某舰面温度的建模及预报

陈永刚, 周晓东, 沈同圣

(烟台海军航空工程学院 控制工程系, 山东 烟台 264001)

**摘 要:** 在系统建模理论的基础上, 运用改进的动态数据建模法, 对南海某舰面温度数据进行了建模, 并通过模型的适用性检验发现, ARMA(6,5)模型是最合适的, 根据建立的模型利用条件期望法进行预报, 结果与原始数据接近得很好。所有的建模和分析过程均在 MATLAB 上实现。

**关键词:** 时间序列; 系统建模法; ARMA 模型; 预报

**中图分类号:** TP273; TP391.9

**文献标识码:** A

**文章编号:** 1671-8798(2005)01-0025-05

## Modeling and forecasting for surface temperature in a Nanhai warship by using system modeling method

CHEN Yong-gang, ZHOU Xiao-dong, SHEN Tong-sheng

(Department of Control Engineering, Yantai Naval Aeronautical Engineering Institute, Yantai 264001, China)

**Abstract:** This paper introduces the dynamic data system modeling method and conditional mathematical expectation predicting method. Using the improved system modeling method, the paper studies the model of the surface temperature of a Nanhai warship and verifies the model's validity. Based on the adaptability test of the model, we can draw the conclusions that the ARMA(6,5) model is most suitable for the surface temperature in a Nanhai warship and the model is stable. Then using the conditional mathematical expectation predicting method, the paper does the simple prediction through the ARMA(6,5) model. The predicting result is rather satisfying. All the modeling and predicting procedures are realized by MATLAB software. The result shows that the methods are simple and easy to be applied to practice.

**Key words:** time series; dynamic data system; ARMA model; predicting

数值气象预报系统理论与技术是 21 世纪初我大气科学的优先发展领域<sup>[1]</sup>。而时间序列分析就讨论各类数据用相应的数学模式去近似描述, 来达

到预测的目的。根据时间序列方法建立的 AR、AR-MA 等模型在气候分析和预测、水文、商情、经济、地震等领域中的应用日趋广泛, 并取得了较好的效果。

收稿日期: 2004-09-02

作者简介: 陈永刚(1979—), 男, 河南孟津人, 硕士研究生, 主要从事智能制导及南海舰面武器装备可靠性研究。

运用何种方案建模,人们已经进行了大量的研究。现有的建模方案主要以自回归滑动平均(ARMA)为主体模型的 Box-Jenkins 方案、Pandit-Wu 方案,另有指数模型、门限回归模型等<sup>[2]</sup>。以 ARMA 为主体模型的 Box-Jenkins 方案是通过预处理实现数据平稳化,再以相关分析来选择模型的形式。这种运算本身可能在最终的序列中引入原来数据所没有的伪周期与趋势。最终拟合的模型虽然是统计上合适的,但它可能给出原来序列结构的一个完全失真的模型。Pandit-Wu 方案是先对数据建立 ARMA( $n, n-1$ )模型,再通过对模型特性的分析,识别序列中的趋势和周期性,这样又使建模过程计算复杂而且阶数较高。其他如门限回归模型(Threshold Auto-Regressive Model, 简称 TAR)、遗传门限自回归模型<sup>[3]</sup>等模型能够有效地描述具有复杂现象的非线性动态系统,是目前时序分析中经常应用的非线性时序模型,但其参数辨识比较复杂,工程应用困难<sup>[4]</sup>。

预报,就是要利用已知时间序列模型所提供的信息来推断一个给定时间序列的未来值。在实际气象数据时间序列分析中,当找到一定的描述模型和参数,获得较为符合实际的拟合效果后,首要的应用就是依据模型作预报<sup>[5]</sup>。目前的各种预测方法都是最小方差意义下的预报,主要的预报方法有 ARMA 序列预报方法, TAR 门限自回归模型预测<sup>[3]</sup>, GM 灰色序列预测<sup>[6]</sup>, 数忆模式预测<sup>[2]</sup>等。本文在前人的既有成果上,简要介绍了动态数据建模法,利用 Pandit-Wu 建模方案,对某年南海某舰舰面温度数据进行了建模研究并利用基于 ARMA 的条件期望法进行了预报。所有的仿真过程均在 MATLAB 软件上实现。

## 1 系统建模方法及其预报理论基础

时间序列分析的一些主要方法都是假定数据样本是来自平稳和各态历经的随机过程,而气象数据序列一般来讲都是非平稳的。由于非平稳过程的复杂性<sup>[5]</sup>,人们尚无完善的处理方法,因而在解决实际气象问题时,总是尽量用差分运算或季节性算子简化数据序列,设法将它们变换成平稳过程。但这种方法如上所述容易引起一系列问题,而使用动态数据系统(DDS, Dynamic Data System)建模方法在拟合模型之前,可以不作差分或季节性算子的预处理,从而可以避免相应的问题。

在建模策略上, DDS 法采用 ARMA( $n, n-1$ )模型去逼近序列 $\{x_t\}$ ,即

$$\begin{aligned} X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_n X_{t-n} = \\ a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_{n-1} a_{t-n+1} \\ a_t \sim N(0, \sigma_a^2) \end{aligned} \quad (1)$$

令  $n$  由小到大以 2 为步长增加,每改变一次阶数,可以根据  $F$ -检验判据判断阶数增加是否有必要。关于  $n$  的取值问题, Pandit-Wu 建议从  $n=2$  开始,按 ARMA( $2n, 2n-1$ )( $n=1, 2, 3, 4, \dots$ )的方式进行<sup>[7]</sup>,具体的叙述及论证参见文献[4]。

预报是系统研究的重要方面,根据系统建模理论我们基于 ARMA 序列进行预报。用 ARMA 序列进行预报的主要方法有格林函数法、逆函数法、条件期望法等,本文利用条件期望法求预报。条件期望的预报可以在正交投影的基础上推导出,也可以用更基本的正交分解概念和三维空间的类比关系来推导。

## 2 建模及预报分析步骤

### 2.1 确定初始猜测值

本文采用基于逆函数的方法来确定初始值。对于 ARMA( $n, n-1$ )模型,首先拟合一 AR( $2n-1$ )模型,即

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_n X_{t-n} + a_t \quad (2)$$

利用式(3)求出 AR( $2n-1$ )模型的系数  $\phi_j$ , ( $j=1, 2, \dots, 2n-1$ ),

$$Y = \begin{bmatrix} X_{2n} \\ X_{2n+1} \\ \vdots \\ X_N \end{bmatrix}, \quad X = \begin{bmatrix} X_{2n-1} & X_{2n-2} & \cdots & X_1 \\ X_{2n} & X_{2n-1} & \cdots & X_2 \\ \vdots & \vdots & \ddots & \vdots \\ X_{N-1} & X_{N-2} & \cdots & X_{N-2n+1} \end{bmatrix},$$

$$\hat{\phi} = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{2n-1} \end{bmatrix}, \text{ 则 } \hat{\phi} = (X'X)^{-1}X'Y \quad (3)$$

其中  $N$  为观测  $X_t$  的个数。

并用它作为 ARMA( $n, n-1$ )模型的前  $2n-1$  个逆函数  $I_j = \phi_j$ , ( $j=1, 2, \dots, 2n-1$ ), 代入(4)式

$$\begin{aligned} (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_{n-1} B^{n-1}) I_j = 0, \\ (j = n+1, n+2, \dots, 2n-1) \end{aligned} \quad (4)$$

其中  $B$  为后移算子:  $B^m I_j = I_{j-m}$ 。

得到  $n-1$  个关于  $\theta$  的线性方程,解方程组可得 ARMA( $n, n-1$ )的滑动平均系数  $\theta_i$ , ( $i=1, 2, \dots$ ),



$n-1$ ),最后由求出的  $\theta_i$  和  $I_j$ ,利用式(5)求出自回归系数:

$$\phi_j = \theta_j - \theta_1 I_{j-1} - \theta_2 I_{j-2} - \cdots - \theta_{j-1} I_1 + I_j \quad (5)$$

其中  $j = 1, 2, \dots, n; \theta_n = 0$ 。

该方法的整个求解过程只涉及到线性方程组的求解,因而计算简便,易于在计算机上编程实现;且与目前其他的参数初值估计方法相比,在参数估计精度上有较大的提高,是对高阶 ARMA 模型进行参数初估计的有效方法。

## 2.2 模型参数终值的估计及模型定阶

由于 ARMA 模型是非线性的,利用非线性最小二乘估计的方法逐步逼近以实现残差平方和的极小化,从而得到合适的参数估计。残差由公式(6)迭代求出。

$$a_t = X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_n X_{t-n} + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_{n-1} a_{t-n+1} \quad (6)$$

采用全局收敛的 Levenberg-Marquardt 修正的高斯-牛顿法算法来迭代计算,得到估计终值后,还可以利用非线性最小二乘理论来求出各参数的置信区间。

求出模型的参数后,可以利用  $F$ -判据,

$$F = \frac{A_1 - A_0}{s} / \frac{A_0}{N-r} \sim F(s, N-r) \quad (7)$$

式(7)中  $A_0$  是不受限模型的平方和(较小),  $A_1$  是受限模型的平方和(较大),  $F(s, N-r)$  是具有  $s$  和  $N-r$  个自由度的  $F$ -分布<sup>[4]</sup>。其中残差平方和的公式为:  $RSS = \sum a_i^2$ 。利用  $F$ -判据可以比较两个模型的平方和以作为停止拟合的判据,也可以进一步用  $F$ -判据判断是否自回归阶次为奇数,而且按置信区间是否包括零,用同样的方式还可以检验滑动平均参数较小的模型的适用性。

在用  $F$ -判据的同时,还要通过残差自相关的  $\chi^2$  分布来检验残差是否是相互独立的,以进一步确定模型的合理性。

## 2.3 利用求得模型进行预报分析

对于任何 ARMA 模型,在时刻  $t$  将要预测的  $X_{t+l}$  的正交分解为

$$X_{t+l} = a_{t+l} + G_1 a_{t+l-1} + G_2 a_{t+l-2} + \cdots$$

其中  $G_j (j = 1, 2, 3, \dots)$  为可以表示系统对任何一个特定  $a_{t+l}$  的动态响应衰减的快慢程度的格林函

数,同时,  $G_j$  的量表示系统记忆冲击  $a_{t+l-j}$  的良好程度。

基于  $a_t, a_{t-1}, a_{t-2}, \dots$  对  $X_{t+l}$  的最好的线性预报为分解中包含  $a_{t-j} (j \geq 0)$  的那一部分,而余下的部分就是预测误差了。这样,

$$X_{t+l} = a_{t+l} + G_1 a_{t+l-1} + \cdots + G_{l-1} a_{t+1} + G_l a_t + G_{l+1} a_{t-1} + \cdots = e_t(l) + \hat{X}(l) = \text{误差} + \text{预报} \quad (8)$$

再由如下的条件期望定则即可求出将来观测的预报<sup>[4,7]</sup>:

$$\begin{aligned} E(X_{t-j}) &= X_{t-j}, j = 0, 1, 2, \dots \\ E(X_{t+j}) &= \hat{X}_t(j), j = 0, 1, 2, \dots \\ E(a_{t-j}) &= a_{t-j}, j = 0, 1, 2, \dots \\ E(a_{t+j}) &= 0, j = 1, 2, 3, \dots \end{aligned} \quad (9)$$

一般 ARMA( $n, m$ ) 模型对预报的 95% 概率极限可用下式<sup>[4]</sup>表示:

$$\hat{X}_t(l) \pm 1.96\sigma_a(1 + G_1^2 + G_2^2 + \cdots + G_{l-1}^2)^{\frac{1}{2}} \quad (10)$$

在  $t+1$  时刻,知道了观测  $X_{t+1}$ ,因而  $a_{t+1} = X_{t+1} - \hat{X}_t(1)$  就成为了已知,在此新信息的基础上,  $X_{t+2}, X_{t+3}, \dots$  等的预报就得到了改善。可以通过下面的更新方程计算出更新的预报:

$$\hat{X}_{t+1}(l) = \hat{X}_t(l+1) + G_l \cdot a_{t+1} \quad (11)$$

更新方程(11)表明,在时刻  $t+1$ ,对观测  $X_{t+1}$  更新后的新预报  $\hat{X}_{t+1}(l)$  可以由它的旧预报  $\hat{X}_t(l+1)$  简单地加上  $G_l$  和新冲击  $a_{t+1}$  的乘积求得。

## 3 建模及预报仿真结果

表 1 是对一年中南海某舰舰面温度数据的模型参数的计算机程序输出,所有过程均在 MATLAB6.5 上实现,其中均值也作为一个参数。

查  $\chi^2$  表知:  $\chi_{0.95}^2(100) = 124.3$ ,查  $F$ -分布表可知:  $F_{0.95}(4, \infty) = 2.37$ ,再与表 1 中的数据相对照,容易看出只有 ARMA(6,5) 模型的  $\chi^2$  值和  $F$ -检验值同时小于查表得出的值,ARMA(6,5) 模型是在 95% 概率水平上最合适的模型,所以,我们不需要再拟合比 ARMA(6,5) 阶次更高的模型了。然后再观察 ARMA(6,5) 模型的参数  $\phi_i$  和  $\theta_i$ ,它们的置信区间都不包括零,所以可以认定 ARMA(6,5) 模型即是最终模型。

表 2 给出了起始点  $t = 90$  开始的 20 步预报及  $X_{90+t}$  的 95% 概率极限和更新预报的计算结果。

表 1 一年中南海某舰舰面温度数据的模型参数

参数	ARMA(2,1)	ARMA(3,2)	ARMA(4,3)	ARMA(5,4)	ARMA(6,5)
$\phi_1$	1.2855 ~ 1.5879	0.35908 ~ 0.85425	- 0.4276 ~ 1.9313	0.93064 ~ 1.2593	0.96118 ~ 1.7172
$\phi_2$	- 0.59063 ~ - 0.30181	0.37986 ~ 1.0211	- 1.0352 ~ 1.3664	- 0.45281 ~ - 0.046321	- 0.84301 ~ - 0.14122
$\phi_3$		- 0.51496 ~ - 0.13346	- 0.91354 ~ 1.5911	- 0.69001 ~ - 0.28313	- 0.35767 ~ 0.13088
$\phi_4$			- 0.73735 ~ 0.19444	0.68896 ~ 1.0762	0.50638 ~ 0.99707
$\phi_5$				- 0.41788 ~ - 0.11063	- 1.2717 ~ - 0.59735
$\phi_6$					0.26098 ~ 0.59087
$\theta_1$	0.73241 ~ 0.93885	- 0.24117 ~ 0.17375	- 1.05 ~ 1.3387	0.38422 ~ 0.62624	0.31912 ~ 1.105
$\theta_2$		0.53356 ~ 0.91061	- 0.62804 ~ 1.0592	- 0.14683 ~ 0.055147	- 0.23977 ~ 0.12634
$\theta_3$			- 0.57002 ~ 1.2699	- 0.73382 ~ - 0.53701	- 0.47251 ~ - 0.14971
$\theta_4$				0.63967 ~ 0.87308	0.61115 ~ 1.0274
$\theta_5$					- 0.83699 ~ - 0.23631
均值 $\mu$	19.017 ~ 26.083	18.633 ~ 26.097	19.849 ~ 24.646	18.831 ~ 26.165	19.194 ~ 25.56
RSS	1386.4	1370.6	1381.1	1295.2	1276
$\chi^2$ 值	126.77	120.66	127.77	116.37	105.97
F 检验	2.0653	5.1659	6.1585	2.5536	1.4372

注:表 1 中所列各模型系数及均值的值为该参数的置信区间,其终值应取置信区间的中值。

表 2 起始点  $t = 90$  开始的 20 步预报及  $X_{90+l}$  的 95% 概率极限和更新预报的计算结果

$l$	$G_l$	$t = 90$	$X_t$	$\hat{X}_t(l)$	$\pm 1.96 \sqrt{Vare_t(l)}$	$\hat{X}_{t+1}(l-1)$
0	1	$t$	20	—	—	—
1	0.62714	$t+1$	24	21.212	$\pm 3.6646$	—
2	0.40448	$t+2$	24	21.411	$\pm 4.3256$	23.159
3	0.43076	$t+3$	24	22.284	$\pm 4.5725$	23.411
4	0.23916	$t+4$	24	22.785	$\pm 4.8374$	23.987
5	0.13599	$t+5$	24	22.667	$\pm 4.9161$	23.334
6	0.15947	$t+6$	24	21.938	$\pm 4.9413$	22.318
7	0.33246	$t+7$	24	21.937	$\pm 4.9758$	22.382
8	0.30083	$t+8$	24	21.954	$\pm 5.1227$	22.881
9	0.28338	$t+9$	24	21.873	$\pm 5.2400$	22.712
10	0.28842	$t+10$	24	21.533	$\pm 5.3419$	22.323
11	0.37149	$t+11$	24	21.746	$\pm 5.4455$	22.550
12	0.30681	$t+12$	24	21.910	$\pm 5.6131$	22.946
13	0.26884	$t+13$	24	21.987	$\pm 5.7246$	22.843
14	0.24704	$t+14$	24	21.813	$\pm 5.8087$	22.563
15	0.29417	$t+15$	25	21.966	$\pm 5.8789$	22.655
16	0.2482	$t+16$	24	22.028	$\pm 5.9769$	22.848
17	0.23322	$t+17$	24	22.050	$\pm 6.0457$	22.742
18	0.22197	$t+18$	24	21.899	$\pm 6.1058$	22.549
19	0.25912	$t+19$	24	21.989	$\pm 6.1598$	22.608
20	0.22822	$t+20$	24	22.011	$\pm 6.2325$	22.734

注:表 2 中  $l$  为预报的步长; $G_l$  为格林函数对应于  $l$  的值; $t$  为预报的起始点; $X_t$  为原始温度数据; $\hat{X}_t(l)$  为以  $t$  为起始点,以  $l$  为预报步长的预报值; $\pm 1.96 \sqrt{Vare_t(l)}$  为预报的 95% 概率极限的范围; $\hat{X}_{t+1}(l-1)$  为更新后的预报值。



图1表示了起始点 $t=90$ 开始的20步预报及 $X_{90+t}$ 的95%概率极限和更新预报。从图上可以看出,更新后的预报和原来的预报都比较接近于原始温度,而更新后的预报较原来的预报得到了改善,利用条件期望法可以较好地进行了预报。

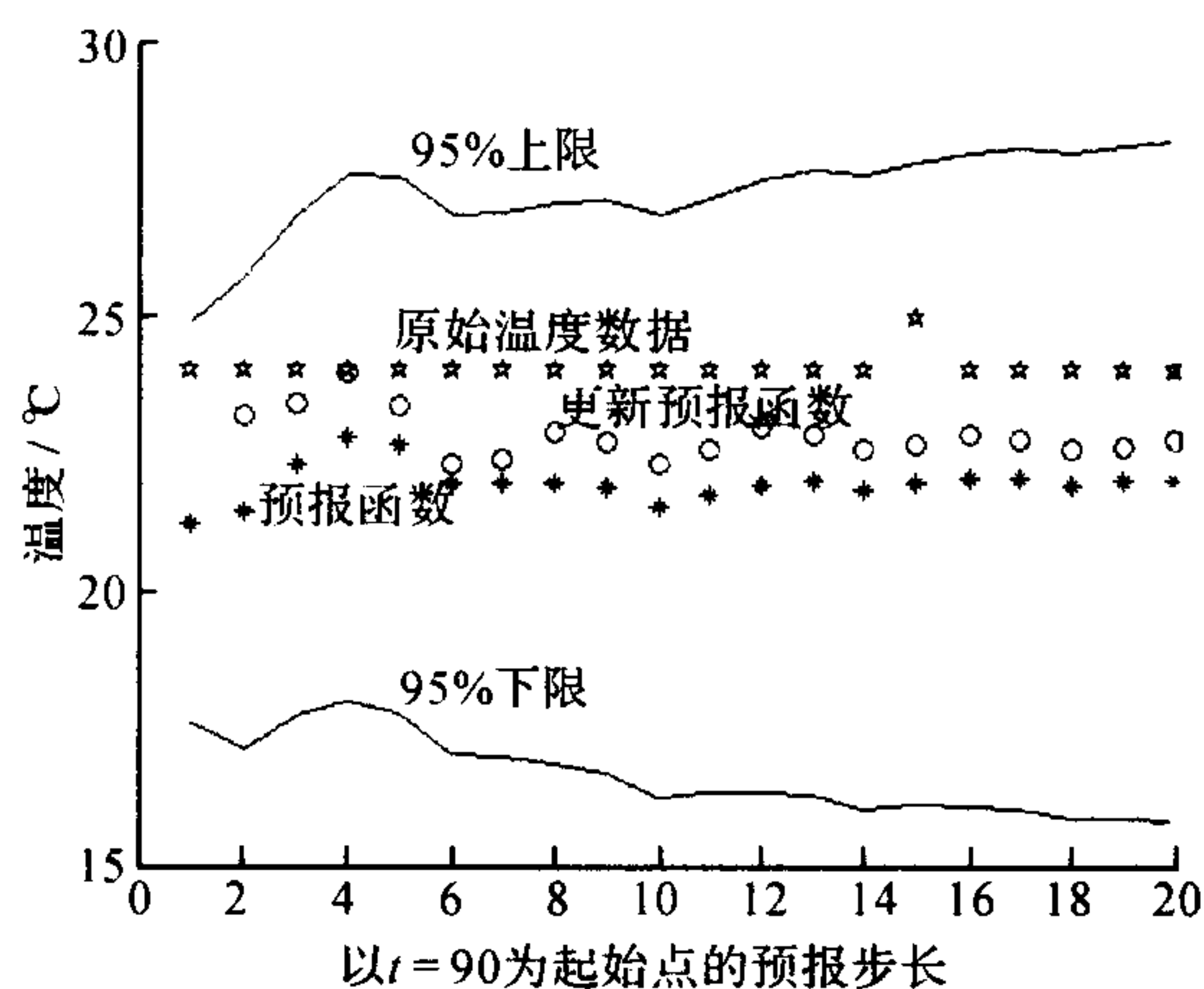


图1 南海某舰舰面温度——ARMA(6,5)模型在起始点 $t=90$ 时的20步预报和更新预报以及95%概率极限

#### 4 结论

通过运用系统建模法对南海某舰舰面温度进行了建模及用条件期望法作了简单的预报,并在MATLAB软件上实现。经过模型的适用性检验发

现,ARMA(6,5)模型比较合适,利用基于ARMA(6,5)模型的条件期望法进行预报,比较接近于原始数据。

整个建模及预报过程可以利用MATLAB软件简单容易的编程实现,从而可以快速准确地确定最终的模型及进行预报,所以,系统建模法是一套简单易行的方案。

#### 参考文献:

- [1] 黄荣辉. 大气科学发展的回顾与展望[J]. 地球科学进展, 2001, 16(5): 643—657.
- [2] 贾晓静, 曹鸿兴. 一种动态数据的新建模法及其预报应用[J]. 应用气象学报, 2002, 13(1): 96—101.
- [3] 金菊良, 杨晓华, 金保明, 等. 遗传门限自回归模型在气象时间序列预测中的应用[J]. 热带气象学报, 2001, 17(4): 415—422.
- [4] 潘迪特 S M, 吴宪民. 时间序列及系统分析与应用[M]. 李昌琪等译. 北京: 机械工业出版社, 1988.
- [5] 丁裕国, 江志红. 气象数据时间序列信号处理[M]. 北京: 气象出版社, 1998.
- [6] 白亚恒, 吴胜军, 黄进良, 等. 基于时间序列的灰色预测技术在估产模型中的应用[J]. 华中师范大学学报(自然科学版), 2001, 35(2): 241—244.
- [7] 杨位钦, 顾 岚. 时间序列分析与动态数据建模[M]. 北京: 北京理工大学出版社, 1988.

## “食品保鲜、加工技术及其产品开发” 通过省级验收

2005年1月8日,由沈莲清教授主持的浙江省科技计划项目“食品保鲜、加工技术及其产品开发”通过了浙江省科技厅组织的项目验收委员会的验收。验收委员会通过听取项目组的汇报、质疑和讨论,一致认为:该项目很好地完成了计划合同内容,下拨科研经费使用合理,为开展更高层次的国际科研合作奠定了基础,一致同意通过项目验收。

生物与化学工程学系