

# 埃尔尼诺的模拟和预测

徐 洪 达

(国家海洋局第一海洋研究所, 青岛)

众所周知, 埃尔尼诺是以出现在美洲秘鲁和厄瓜多尔沿海区域数千千米尺度的异常高的暖水为特征的异常的海洋和天气事件. 当埃尔尼诺现象出现时, 全球性的大气环流发生变化, 出现全球性天气异常. 因而, 它已被看成是全球年际气候变化的最突出的一种标志. 所以, 模拟和预测埃尔尼诺现象, 愈来愈引起海洋学家和气象学家的研究兴趣.

为了诊断研究埃尔尼诺现象, 海洋学家和气象学家已提出了许多类型的埃尔尼诺指数<sup>[5-7]</sup>. 本文试图用Box-Jenkins的建模方法<sup>[4]</sup>, 通过埃尔尼诺指数来模拟和预测埃尔尼诺现象.

## 一、建模方法

Box-Jenkins建模方法, 在数学上比较完整, 预测精度也较高. 用这种方法可以求得平稳时间序列的自回归(AR)模式、自回归滑动平均(ARMA)模式和门限(Threshold)模式等. 这里用Box-Jenkins方法来寻求模拟和预测埃尔尼诺的ARMA模式.

定义: 把埃尔尼诺指数时间序列

$$\{x_t, t=1, 2, \dots\}$$

看成是平稳时间序列, 把 $t$ 时刻项 $X_t$ 表示成 $p$ 个最近的项 $x_{t-1}, x_{t-2}, \dots, x_{t-p}$ 和 $q$ 个最近的预测误差 $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$ 以及即时噪声 $\varepsilon_t$ 的线性函数, 即把 $\{x_t\}$ 表示为

$$x_t = \theta_0 + \sum_{h=1}^p \phi_h x_{t-h} + \varepsilon_t - \sum_{h=1}^q \theta_h \varepsilon_{t-h}. \quad (1)$$

其中,  $\{\varepsilon_t\}$ 是均值为零、方差为 $\delta^2$ 的白噪声序列, 则式(1)叫做埃尔尼诺指数序列的 $(p, q)$ 阶自回归滑动平均模式, 简记为ARMA $(p, q)$ . 式(1)中的系数 $\phi_1, \phi_2, \dots, \phi_p$ 称为自回归系数,  $\theta_0, \theta_1, \theta_2, \dots, \theta_q$ 称为滑动平均系数.

由式(1)可见, 当阶数 $(p, q)$ 确定时, 要得到 $\{x_t\}$ 的自回归滑动平均模式, 关键问题是确定自回归系数和滑动平均系数. 计算这两种系数可用多种方法, 这里用如下方法, 其主要计算步骤如下.

### 1. 计算 $\{x_t\}$ 的自相关函数估计值

$$\widehat{R}_k = \frac{1}{N} \sum_{t=1}^{N-k} (x_t - \bar{x})(x_{t+k} - \bar{x}), k=0, 1, 2, \dots, p+q. \quad (2)$$

其中,

$$\bar{x} = \frac{1}{N} \sum_{t=1}^N x_t,$$

## 2. 解 $p$ 阶线性方程组

$$\begin{pmatrix} \widehat{R}_q & \widehat{R}_{q-1} & \cdots & \widehat{R}_{q-p+1} \\ \widehat{R}_{q+1} & \widehat{R}_q & \cdots & \widehat{R}_{q-p+2} \\ \vdots & \vdots & & \vdots \\ \widehat{R}_{q+p-1} & \widehat{R}_{q+p-2} & \cdots & \widehat{R}_q \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{pmatrix} = \begin{pmatrix} \widehat{R}_{q+1} \\ \widehat{R}_{q+2} \\ \vdots \\ \widehat{R}_{q+p} \end{pmatrix} \quad (3)$$

其解记为  $\widehat{\phi}_1, \widehat{\phi}_2, \dots, \widehat{\phi}_p$ , 此解即为自回归系数的估计值.

## 3. 按下式计算 $\theta_0$ 的估计值,

$$\theta_0 = \begin{cases} \bar{x}, & \text{当 } p=0, \\ \bar{x} \left( 1 - \sum_{i=1}^p \widehat{\phi}_i \right), & \text{当 } p \neq 0. \end{cases}$$

## 4. 取 $\widehat{\phi}_0 = -1$ , 计算

$$R'_k = \sum_{i,j=0}^p \widehat{\phi}_i \widehat{\phi}_j \widehat{R}_{|k+i-j|}, \quad k=0, 1, 2, \dots, q,$$

然后取  $\tau_0^{(0)} = \sqrt{R'_0}$ ,  $\tau_1^{(0)} = \tau_2^{(0)} = \dots = \tau_q^{(0)} = 0$ , 用线性化迭代法解如下方程组

$$\begin{cases} \tau_0^2 + \tau_1^2 + \dots + \tau_q^2 = R'_0, \\ \tau_0 \tau_1 + \tau_1 \tau_2 + \dots + \tau_{q-1} \tau_q = R'_1, \\ \tau_0 \tau_2 + \tau_1 \tau_3 + \dots + \tau_{q-2} \tau_q = R'_2, \\ \dots\dots \\ \tau_0 \tau_q = R'_q. \end{cases} \quad (4)$$

设  $\tau_i^{(m)}$  ( $i=0, 1, 2, \dots, q$ ) 为经过  $m$  次迭代所得到的满足预先给定精度的方程组 (4) 的解, 则滑动平均系数  $\theta_i$  ( $i=1, 2, \dots, q$ ) 的估计值为

$$\widehat{\theta}_i = -\frac{\tau_i^{(m)}}{\tau_0^{(m)}}, \quad i=1, 2, \dots, q. \quad (5)$$

模式的阶数 ( $p, q$ ) 用赤池 (Akaike) 准则<sup>(3)</sup> 确定. 即先给定模式的极大阶数为  $L$ , 对  $p, q=1, 2, \dots, L$  计算

$$\text{AIC}(p, q) = N \times \text{Log} \delta_{p,q}^2 + 2(p+q+2), \quad (6)$$

然后取

$$\text{AIC} = \min\{\text{AIC}(p, q) : p, q=1, 2, \dots, L\},$$

这个 AIC 所对应的阶数 ( $p, q$ ) 即为所求模式的阶数. 将这个阶数记为  $(p_0, q_0)$ , 与此阶数对应的系数记为  $\widehat{\phi}_1, \widehat{\phi}_2, \dots, \widehat{\phi}_{p_0}, \widehat{\theta}_0, \widehat{\theta}_1, \dots, \widehat{\theta}_{q_0}$ , 这就是所求的自回归系数和滑动平均系数.

白噪声方差  $\delta_{p,q}^2$  按下式估计:

$$\hat{\delta}_{p,q}^2 = \begin{cases} -\sum_{k=0}^p \hat{\phi}_k \hat{R}_k, & \text{当 } q=0, \\ [\tau_0^{(m)}]^2, & \text{当 } q \neq 0. \end{cases} \quad (7)$$

阶数和系数确定之后, 所求模式就是

$$x_t = \hat{\theta}_0 \sum_{k=1}^{p_0} \hat{\phi}_k x_{t-k} + \varepsilon_t - \sum_{k=1}^{q_0} \hat{\theta}_k \varepsilon_{t-k}. \quad (8)$$

有了这个模式就可对埃尔尼诺进行模拟和预测.

## 二、所用资料

本文选用如下两种埃尔尼诺指数的12个月滑动平均滤波的时间序列来模拟.

### 1. SSTI (海表面温度指数) <sup>(7)</sup>

此指数序列是1950年1月至1983年12月期间, 中赤道太平洋和东赤道太平洋区域 ( $6^\circ\text{N}$ — $6^\circ\text{S}$ ,  $90^\circ$ — $180^\circ\text{W}$ ) 所有  $4^\circ$ (纬度)  $\times$   $4^\circ$ (经度) 方区的逐月海表面水温距平的平均值. 其中1950—1979年的数据是Wright (1984) 根据美国环境研究实验室的Fletcher所整理的1954—1979年期间上述区域中  $4^\circ \times 10^\circ$  方区逐月海表面温度距平值重新计算得到的. 1980—1983年的数据是美国气候分析中心的Reynolds整理的.

### 2. ENI <sup>(6)</sup>

此指数序列是Weare等 (1984) 根据1949—1973年  $20^\circ\text{S}$  以北太平洋区域海表面温度序列用经验正交函数分析法得到的经验正交函数而确定的时间系数和延伸的时间系数. 其中, 1949—1973年部分是按下式计算得到的.

$$\text{ENI}_i = \sum_{j=1}^{160} \text{EOF}_j \times \text{SST}_{i,j}.$$

式中  $i$  是1974年1月至1983年12月之间的月标号,  $j$  是上述经验正交函数 (EOF) 图上有海表面温度距平的160个空间点的标号.

## 三、模式和结果

众所周知, 1982—1983年期间曾出现过一次较强的埃尔尼诺事件. 因此, 为了较好地检验用 ARMA ( $p_0, q_0$ ) 模式模拟和预测埃尔尼诺的效果, 在建模时选用 SSTI 和 ENI 序列的时限至1982年7月. 这两个指数序列的12个月滑动平均滤波结果如图1所示. 图中实线表示1982年7月之前的部分, 带  $\times$  号虚线表示1982年7月之后的部分.

计算结果如下.

### 1. SSTI 序列

阶数:  $(p_0, q_0) = (5, 0)$ ;

自回归系数:  $\hat{\phi}_1 = -1.2487, \hat{\phi}_2 = -0.0971, \hat{\phi}_3 = -0.0397, \hat{\phi}_4 = -0.0528, \hat{\phi}_5 = -0.1173$ .

滑动平均系数:  $\hat{\theta}_0 = -0.0005$ .

噪声方差  $V_{\varepsilon\varepsilon} = 0.0087$ .

由此得12个月滑动平均滤波的SSTI的ARMA(5, 0)模式为

$$x_t = -0.0005 + 1.2487x_{t-1} - 0.0971x_{t-2} - 0.0397x_{t-3} - 0.0528x_{t-4} - 0.1173x_{t-5} + \varepsilon_t \quad (9)$$

## 2. ENI序列

阶数:  $(p_0, q_0) = (2, 1)$ ;

自回归系数:  $\hat{\phi}_1 = 1.9496, \hat{\phi}_2 = -0.9701$ ;

滑动平均系数:  $\hat{\theta}_0 = -0.0020, \hat{\theta}_1 = 0.7728$ ;

噪声方差  $V_{\varepsilon_t} = 0.2245$ .

由此得12个月滑动平均滤波的ENI的ARMA(2, 1)模式为

$$x_t = -0.0020 + 1.9496x_{t-1} - 0.9701x_{t-2} + \varepsilon_t - 0.7728\varepsilon_{t-1} \quad (10)$$

用模式(9)和(10)模拟和预测的结果如图1中的虚线所示。

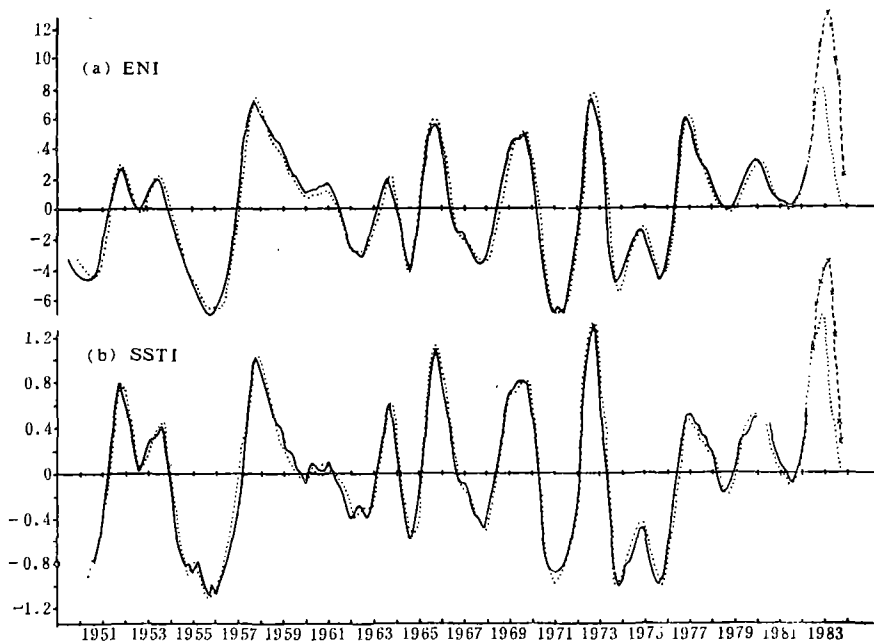


图1 12个月滑动平均的、模拟和预测的埃尔尼诺指数序列

由图1明显可见:

(1) 模拟得到的埃尔尼诺指数序列与原序列吻合得相当好, 也就是说模拟出了1950—1982年之间发生的8次埃尔尼诺事件. 它们分别发生在1951, 1953, 1957—1958, 1963, 1965, 1969, 1972, 1976—1977年;

(2) 预测出的SSTI和ENI时间序列, 于1982—1983年间呈现出正的较大的距平值, 这表明此期间出现较强的埃尔尼诺事件, 显然与实际情况是一致的(但就数值上来讲, 预测出的埃尔尼诺指数值较实测值小).

综上所述, 可以看到12个月滑动平均的SSTI的ARMA(5, 0)模式和12个月滑动平

均的ENI的ARMA(2, 1)模式均能较好地模拟和预测出埃尔尼诺事件。因此,我们可以用12个月滑动平均滤波的埃尔尼诺指数的ARMA( $p, q$ )模式来模拟和监测埃尔尼诺现象。

### 参 考 文 献

- [1] 中国科学院计算中心概率统计组, 概率统计计算, 科学出版社, 1979, 267—280.
- [2] 郑宗成等, 实用预测方法BASIC程序库, 中山大学出版社, 1985, 229—246.
- [3] Akaike, H., Information theory and an extension of the maximum likelihood principle, *Second International Symposium on Information Theory*, 1973, 267—281.
- [4] Box, G.E.P. and G.M. Jenkins, *Time Series Analysis, Forecasting and Control*, San Francisco, Holden-Day, 1970.
- [5] Rasmusson, E.M. et al., Meteorological aspects of the El Nino/Southern Oscillation, *Monthly Wea. Rev.*, 112 (1983), 184—188.
- [6] Weare, B.C., An extension of an El Nino index, *Monthly Weather Review*, 114 (1985), 644—647.
- [7] Wright, P.B., Relationship between indices of the Southern Oscillation, *Monthly Weather Review*, 113 (1984), 1913—1919.