第13卷 第3期 郑州工学院学报 Vol.13 No.3
1992年 9 月 Journal of Zhengzhou Institute of Technology Sep. 1992

月径流序列的季节性 ARIMA 模型

贺北方 吕廷军* *

(郑州工学院水环系)

摘 要: 基于月径流序列是一类周期性的非平稳时间序列的特点，本文建立了季节性 *ARIMA* 模型. 文中编制了电算程序，并应用于水库的月径流预报，获得了满意的结果.

关键词: 月径流，时间序列分析，预报。

中国图书分类号: O213

1 月径流序列的季节性模型

月径流序列呈现以年为周期的季节性变化，其统计特性随季节而异，是一种季节性的非平稳时间序列. 对于这类非平稳序列，各种平稳时间序列模型不能直接应用. 因此，有必要探讨这种非平稳时间序列建模及预报方法问题.

1.1 季节性自回归模型

月径流序列$\{x_{ij}\}$，可通过标准化处理以消除均值、方差的季节性变化影响:

$$z_{ij}=(x_{ij}-\mu_j)/\sigma_j$$

式中 z_{ij} 为标准化变量序列. i=1, 2, …n, n 为年数; j=1, 2, …12 为月份. μ_j 和 σ_j 分别为第 J 月的均值和方差. 对标准化的月径流序列 Z_{ij} 可建立如下的自回归模型 AR(P):

$$Z_{i,j}=\varphi_{1,j}Z_{i,j-1}+\varphi_{2,j}Z_{i,j-2}+\cdots+\varphi_{p,j}Z_{i,j-p}+\varepsilon_{i,j}$$

该模型的特点是自回归系数 φ_j 随月份而变. 这种变参数的自回归模型，称为季节性自回归模型.

1.2 季节性 ARIMA 模型

* 收稿日期: 1990-11-15

* * 张红军参加了资料分析及模型检验工作.

月径流序列的季节性变化所引起的非平稳性，可以通过季节差分消除。一般，若ω为周期（对月径流序列$\omega=12$），∇为差分算子，B为后移算子，则

$$\nabla_\omega X_t=(1-B^\omega)X_t=X_t-X_{t-\omega}$$

令$u_t=X_t-X_{t-12}$，为月度差分后的新序列．对u_t建立自回归移动平均模型ARMA，但对季节性的原序列X_t而言，乃建立了季节性的积分自回归移动平均模型，记为季节性ARIMA模型:

$$\varphi(B)(1-B^\omega)^D X_t=\theta(B)\varepsilon_t$$

式中D为差分次数，ε_t为随机干扰．φ（B）和θ'（B）分别称为自回归算子和移动平均算子:

$$\varphi(B)=(1-\varphi_1B-\varphi_2B^2-\cdots-\varphi_pB^p)$$

$$\theta(B)=(1-\theta_1B-\theta_2B^2-\cdots-\theta_qB^q)$$

1.3 乘积型的季节性ARIMA模型

乘积型的季节性ARIMA模型，可以表示复杂的季节性时间序列的各种统计特性，但模型的识别、参数的估计、检验等比较复杂。一般的乘积型ARIMA(p,d,q)×(P,D,Q)$_\omega$模型可表示为:

$$\Phi(B^\omega)\varphi(B)(1-B^\omega)^D(1-B)^d x_t=\Theta(B^\omega)\theta(B)\varepsilon_t$$

式中: $$\Phi(B^\omega)=(1-\varphi_1B^\omega-\varphi_2B^{2\omega}-\cdots-\varphi_PB^{P\omega})$$

$$\Theta(B^\omega)=(1-\theta_1B^\omega-\theta_2B^{2\omega}-\cdots-\theta_QB^{Q\omega})$$

因该模型系由ARIMA(p，d，q)与$ARIMA(P,D,Q)_\omega$通过两个多项式相乘而得，故称为乘积型季节性ARIMA模型．式中d，D分别为一般差分与季节差分的阶次．

2 鲇鱼山水库月径流序列模型的建立

对鲇鱼山水库，我们已编制了考虑面临时段径流预报的优化调度方案．要实施水库优化调度，应配置相应的月（旬）径流序列预报方案．

用时间序列分析方法进行月径流序列预报，是将月径流序列$\{x_t, t\in T\}$视为随机过程，以样本序列$\{X_t\}$，t=1，2，…N建立随机序列的统计模型，然后用以往和现在的资料对未来径流过程作出预报．其建模过程如图1所示．

2.1 确定基本模型形式

鲇鱼山水库位于淮河支流灌河上，坝址有1932--1987年56年径流系列．现采用1953--1987年的35年实测月径流序列$\{X_t\}$作为建模的原始序列．

鉴于月径流序列X_t具有明显的以年为周期的季节性变化，是一种季节性的非平稳时间序列．因此，采用季节性模型作为基本模型．

2.2 模型识别

模型识别，就是从基本模型簇中选择一个与预测对象的实际过程相吻合的模型，即确定模型的类型和阶数．

模型识别的方法应以自相关函数 ρ (k)、偏自相关函数 φ_{kk} 识别为主，辅以直观检验。

样本自相关函数：

$$\hat{\rho}(k)=\frac{\hat{R}(k)}{\hat{R}(0)}$$

式中：
$$\hat{R}(k)=\frac{1}{N-K}\sum_{t=1}^{N-K}(X_t-\bar{X})(X_{t+k}-\bar{X})$$

$$\hat{R}(0)=\sigma^2=\frac{1}{N-K}\sum_{t=1}^{N-K}(X_t-\bar{X})^2$$

$$\bar{X}=\frac{1}{N}\sum_{t=1}^{N}X_t \quad (N\text{为样本容量})$$

样本偏自相关函数 φ_{kk}，可用下述递推算法：

$$\hat{\varphi}_{1,1}=\hat{\rho}(1)$$

$$\hat{\varphi}_{m,m}=[\hat{\rho}(m)-\sum_{j=1}^{m-1}\hat{\varphi}_{m-1,j}\hat{\rho}(m-j)][1-\sum_{j=1}^{m-1}\hat{\varphi}_{m-1,j}\hat{\rho}(j)]^{-1}$$

$$(m=2,\ 3,\ 4,\ \cdots\cdots)$$

$$\hat{\varphi}_{m,j}=\hat{\varphi}_{m-1,j}-\hat{\varphi}_{m,m}\cdot\hat{\varphi}_{m-1,m-j}$$

$$(j=1,\ 2,\ 3,\ \cdots\cdots,\ m-1)$$

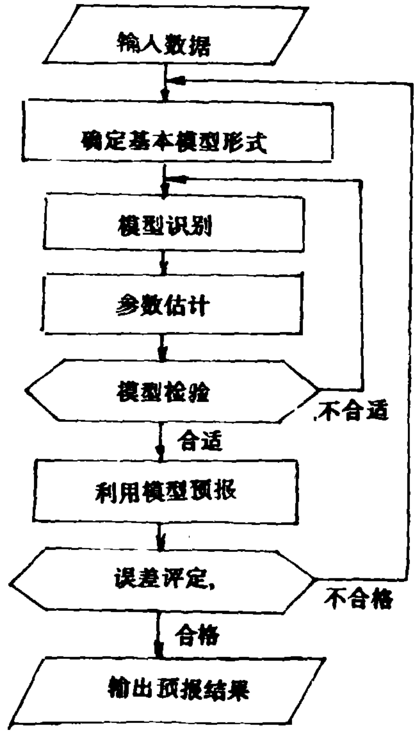

图 1　建模流程图

不同类型的模型具有各自不同的自相关函数和偏自相关函数特性，这使我们有可能利用样本序列的 $\hat{\rho}(k)$ 和 $\hat{\varphi}_{kk}$ 进行模型识别。

表 1　时间序列模型识别条件

模　　型	自相关函数 ρ(k)	偏自相关函数 φ_{kk}
自回归模型 AR(p)	指数衰减，拖尾	K＞p 时 $\varphi_{kk}=0$，截尾
移动平均模型 MA(q)	K＞q 时，$\rho(k)=0$，截尾	指数衰减，拖尾
自回归移动平均模型 ARMA(p,q)	拖尾，无截止点	拖尾，无截止点

对于鲇鱼山水库的月径流序列 X_t，先进行了中心化处理：$y_t=x_t-\bar{x}$；然后对 y_t 的样本序列计算自相关函数 $\hat{\rho}(k)$ 和偏自相关函数 $\hat{\varphi}_{kk}$，因 $\hat{\rho}(k)$ 在 k＝12，24，……处出现高峰，说明序列 y_t 存在 $\omega=12$ 的周期波动，因此，采用月度差分以消除之。即

$$V_t=\nabla_{12}y_t=(1-B^{12})y_t$$

经月度差分后的 V_t 序列仍不平稳，该序列的 $\hat{\rho}(k)$ 和 $\hat{\varphi}_{kk}$ 衰减非常缓慢，可能是 ARIMA（p，d，q）模型。再对 V_t 序列进行一阶差分：

$$U_t=\nabla V_t=(1-B)V_t$$

U_t 序列已趋平稳。从 U_t 序列的样本自相关函数 $\hat{\rho}(k)$ 和样本偏自相关函数 $\hat{\varphi}_{kk}$ 看，两

者均无截止点，但呈几何衰减，说明 U_t 序列可能是 ARMA(p,q)模型．其中阶次 p，q 的确定，尚待从低阶到高阶而逐步进行参数估计、模型检验的尝试，直至模型被接受为止．

2.3 参数估计

经上述模型识别，可以建立 U_t 序列的 ARMA(p,q)模型：

$$U_t-\varphi_1U_{t-1}-\varphi_2U_{t-2}-\cdots-\varphi_pU_{t-p}=\varepsilon_t-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2}\cdots-\theta_q\varepsilon_{t-q}$$

若写成算子形式，则

$$\Phi(B)U_t=\theta(B)\varepsilon_t$$

式中 φ_1，φ_2，……，φ_p 和 θ_1，θ_2，……，θ_q 分别称为自回归系数及移动平均系数，统称为模型参数．这些参数需根据样本序列进行估计．参数估计常用的方法有矩法、非线性最小二乘法、广义最小二乘法等，可根据模型类型及精度要求采用相应的方法．

对鲇鱼山水库的月径流序列 U_t，ARMA 模型的参数初估计采用了矩法，参数的精估计采用了非线性最小二乘法．

2.4 模型检验

若选用的模型是合适的，则用此模型所描述的过程与实际过程间的残差，应当是白噪声序列．

残差 $e_t=\hat{U}_t-U_t$

残差 e_t 的样本自相关函数为

$$\hat{\rho}_k(e_t)=\sum_{t=1}^{N-K}(e_te_{t+k})\Big/\sum_{t=1}^{N}e_t^2$$

构造统计量 Q

$$Q=N\sum_{k=1}^{m}\hat{\rho}_k^2(e_t)$$

Q 渐近服从自由度为(m−p−q)的 X^2 分布．式中 N 为样本容量，m 为最大滞时，一般取$m\leqslant\frac{N}{4}$．

若 $$Q\leqslant X_\alpha^2(m-p-q)$$

可认为模型选得合适．否则，应对模型进一步识别或重新构造模型．式中 α 为显著性水平，α 常取 0.05 或 0.01；$X_\alpha^2(m-p-q)$是由X^2 分布表查得的临界值．

鲇鱼山水库的月径流序列 U_t，初拟为 ARMA（p，q）模型．取 $\alpha=0.05$，m=30，由低阶到高阶试验选定阶次 p 和 q，按上述方法进行模型检验，最后选定 p=2，q=0，即对 U_t 序列选用了二阶自回归模型 AR(2):

$$U_t=\varphi_1U_{t-1}+\varphi_2U_{t-2}+\varepsilon_t$$

而对中心化的原序列 y_t 而言，则为季节性的 ARIMA 模型：

$$\varphi(B)\nabla\nabla_{12}y_t=\varepsilon_t$$

即 $$(1-\varphi_1B-\varphi_2B^2)(1-B)(1-B^{12})y_t=\varepsilon_t$$

此模型既利用了以年为采样周期的相关信息，又利用了以月为采样周期的相关信息，

描述了含有线性趋势和月度周期波动的随机时间序列．式中自回归系数估计为：$\varphi_1=-0.626$，$\varphi_2=-0.320$；满足AR（2）序列的平稳性条件．

3 月径流序列预报

3.1 月径流序列预报公式

如上所述，鲇鱼山水库的月径流序列选用了季节性ARIMA（2，1，0）$_{12}$模型，即

$$(1-\varphi_1 B-\varphi_2 B^2)(1-B)(1-B^{12})y_t=\varepsilon_t$$

将$\varphi_1=-0.626$，$\varphi_2=-0.320$代入并整理得：

$$y_t=0.374y_{t-1}+0.306y_{t-2}+0.32y_{t-3}+y_{t-12}-0.374y_{t-13}-0.306y_{t-14}-0.32y_{t-15}+\varepsilon_t$$

以(K+l)代替t，对上述模型的等式两边取数学期望，可得K时刻外推l步的月径流预报公式为

$$\hat{y}_k(l)=E[y_{k+l}]=0.374\hat{y}_{k+l-1}+0.306\hat{y}_{k+l-2}+0.32\hat{y}_{k+l-3}+y_{k+l-12}-0.374y_{k+l-13}-0.306y_{k+l-14}-0.32y_{k+l-15}$$

不同步长l的预报公式为：

$l=1$:
$$\hat{y}_k(1)=0.374y_k+0.306y_{k-1}+0.32y_{k-2}+y_{k-11}-0.374y_{k-12}-0.306y_{k-13}-0.32y_{k-14}$$

$l=2$:
$$\hat{y}_k(2)=0.374\hat{y}_k(1)+0.306y_k+0.32y_{k-1}+y_{k-10}-0.374y_{k-11}-0.306y_{k-12}-0.32y_{k-13}$$

⋮

⋮

$l=12$:
$$\hat{y}_k(12)=0.374\hat{y}_k(11)+0.306\hat{y}_k(10)+0.32\hat{y}_k(9)+y_k-0.374y_{k-1}-0.306y_{k-2}-0.32y_{k-3}$$

利用上述递推预报方法，可在年底预报出下一年12个月的月径流量$\hat{X}_k(l)$:

$$\hat{X}_k(l)=\hat{y}_k(l)+\bar{X}\qquad(l=1,2,3,\cdots 12)$$

3.2 新息实时校正预报

在水库月径流预报过程中，随着时间的推移，可以不断地获得月径流的实测值，从而可以利用这种新的信息校正原来的预报值，提高预报精度．

利用新息改善预报精度有两种途径：其一是适时地根据量测数据和估计结果，自行调整模型参数．称为参数的在线辨识，使参数具有“跟踪”能力，以适应非平稳或统计特性“时变”的序列．常用的实时算法有递推最小二乘法和自适应卡尔曼滤波法．其二是不修改模型参数，仅利用新的信息进行现时改正预报：

$$\hat{y}_{k+1}(l)=\hat{y}_k(l+1)+G_l[y_{k+1}-\hat{y}_k(1)]$$

式中y_{k+l}是过程到了$(k+l)$时刻所得到的月径流实测值，$\hat{y}_k(1)$是立足于k时刻的一步

预报值，两者之差：$y_{k+1}-\hat{y}_k(1)=a_{k+1}$，是$k$时刻的一步预报误差，称为“新息”．上式表明，利用新息a_{k+1}可以对k时刻所作的$(l+1)$步预报$\hat{y}_k(l+1)$进行校正，校正后的预报值$\hat{y}_{k+1}(l)$是原有预报值$\hat{y}_k(l+1)$与新息a_{k+1}的加权和．式中G_l为权重，可用下式求得：

$$G_j=\varphi_1 G_{j-1}+\varphi_2 G_{j-2}+\cdots\cdots+\varphi_p G_{j-p}-\theta_j, \quad (j\geqslant 1)$$

其中 $\theta_j=0$，当$j>q$时

$$G_j=\begin{cases}0, & 当 j<0 时\\ 1, & 当 j=0 时\end{cases}$$

对鲇鱼山水库月径流预报而言，$G_0=1$；$G_1=\varphi_1=-0.626$；$G_j=-0.626G_{j-1}-0.32G_{j-2}$，　$(j\geqslant 2)$．

在月径流预报过程中，首先立足于K时刻作预报，当过程值y_{k+1}已知后，可把预报立足点向前移动一步，作（k+1）时刻的l步预报：

$$\hat{y}_{k+1}(l)=\hat{y}_k(l+1)+G_l[y_{k+1}-\hat{y}_k(1)]$$

这种校正预报因充分利用了新的实测信息，且预报步程缩短，使预报精度得以改善和提高。这种方法不需要改变模型参数，简便易行，适宜于较长步程的预报．

4 模型实用性分析

选定模型是否反映了随机序列真实的统计特性，需对预报误差进行分析，以评定和检验随机模型预报的可靠性和预报精度．模型的评定和检验，均是分析预报误差情况，确定模型的有效性．区别在于：评定用建模的全部序列，检验则用预留序列．评定和检验的定量指标一般用合格率，平均偏移B，确定性系数R^2等．

平均偏移

$$B(l,j)=\frac{1}{n}\sum_{i=1}^{n}[x_{i,j}-\hat{x}_{i,j}(l)]$$

确定性系数

$$R^2(l,j)=1-\frac{MSE(l,j)}{S_j^2}$$

其中：$$MSE(l,j)=\frac{1}{n}\sum_{i=1}^{n}[x_{i,j}-\hat{x}_{i,j}(l)]^2$$

$$S_j^2=\frac{1}{n}\sum_{i=1}^{n}(x_{i,j}-\hat{x}_j)^2 \qquad (i=1,2,\cdots n年；j=1,2,\cdots 12月)$$

式中MSE为预报值的均方差；S_j^2为实测水文序列$x_{i,j}$第j月的方差．显然，若B值愈接近于零，R^2愈接近于1，预报精度愈高，所选用的模型愈好．据水电部《水文情报规范》规定：$R^2>0.90$，为甲等方案，$R^2=0.70\sim0.89$为乙等，$R^2=0.50\sim0.69$为丙等．

据此对鲇鱼山水库月径流预报模型进行了评定，计算结果如表2.

表2　鲇鱼山水库月径流预报模型评定表

月份	1	2	3	4	5	6	7	8	9	10	11	12
R^2	0.99	0.98	0.98	0.93	0.92	0.82	0.80	0.78	0.93	0.92	0.98	0.99
等级	甲	甲	甲	甲	甲	乙	乙	乙	甲	甲	甲	甲

由表2看出，除汛期6～8月的 R^2 在0.80左右而属乙等外，其余月份的 R^2 均在0.92以上，均属甲等。说明预报精度尚好，所建的鲇鱼山水库月径流预报模型是有效的。

该模型曾用于鲇鱼山水库近几年的月径流预报。1987—1990年间正值该水库月径流序列丰枯交替时期，1987年出现了年水量达 $1112\times10^6m^3$ 的丰水年，相继而来的1988年却是年水量为 $274\times10^6m^3$ 的枯水年份。径流序列的这种丰枯转折，给预报带来了一些困难，使个别时段的径流预报不太准确，但总的说来，径流预报的精度还是较好的。尤其是采用了逐时段实时校正预报，使月径流预报精度明显提高，可以满足优化调度所提出的面临时段径流预报的要求，提高了水库优化调度的预见性和主动性。

参考文献

〔1〕 杨位钦等. 时间序列分析与动态数据建模. 北京理工大学出版社，1988年

〔2〕 丁晶等. 随机水文学. 成都科技大学出版社，1988年

〔3〕 M.B. Priestley. Spectral Analysis and Time Series. London，1981

A Seasonal ARIMA Model of Monthly Runoff Series

He Beifang Lu Tingjun

(Zhengzhou Institute of Technology)

Abstract: The monthly runoff series is a kind of the seasonal non-stationary time series. A seasonal ARIMA model of monthly runoff series is proposed in this paper, it is in accordance with the special charateristic of the monthly runoff. It's algorithm is programmed in FORTRAN language and is used in the forecast of monthly runoff of resourvoir, the result is satisfactory.

Keywords: Monthly runoff, time series analysis, forecast.