II－8　時系列モデルによる年水文量のシミュレーション

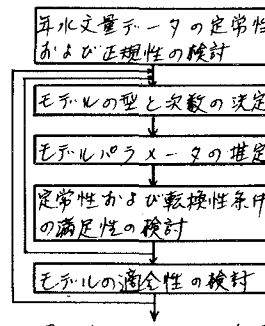
群馬高専　正員　山本好克

1.はじめに　水資源計画において有用となろう水文量の模擬発生あるいは将来予測などのいわゆるシミュレーションについて，定常時系列モデルである自己回帰（AR）型，移動平均（MA）型および混合（ARMA）型モデルにて考察し，適用例によりこれらモデルによるシミュレーション手法の有用性について検討する。

2.モデル式およびシミュレーション手法　N年間の年水文量データ $y_1, y_2, \cdots, y_N$ を用いて，AR(p)，MA(q) あるいは ARMA(p, q) 型にモデル化する手順は，図－1のごとく設定できよう[1]。この手順により確立されるARMA(p, q) モデルの一般式は，次式となる。

$y_t = \hat{\mu} + \hat{\sigma} z_t$ ----(1)　　$z_t = \sum_{j=1}^{p} \hat{\phi}_j z_{t-j} + \varepsilon_t - \sum_{j=1}^{q} \hat{\theta}_j \varepsilon_{t-j}$ ---(2)

ここで，$\hat{\mu}$，$\hat{\sigma}$ は，y_t ($t=1, \cdots N$) の平均と分散，$\hat{\phi}_1, \cdots, \hat{\phi}_p$ は，自己回帰係数，$\hat{\theta}_1, \cdots, \hat{\theta}_q$ は，移動平均係数，ε_t は，平均0，分散 $\hat{\sigma}_\varepsilon^2$ の独立正規ランダム量である。



図－1　モデル化の手順

式(2)において，$q=0$ であればAR(p)モデルに，$p=0$ であればMA(q)モデルとなるので，ARMA(p, q)モデルの一般式を用いた年水文量のシミュレーション手法が考察できる。

(1) 模擬発生手法　式(2)の ε_t を，$\eta_t = \varepsilon_t / \hat{\sigma}_\varepsilon$ の標準正規ランダム変量を導入して書き表わすと次式となる。　$z_t = \sum_{j=1}^{p} \hat{\phi}_j z_{t-j} + \hat{\sigma}_\varepsilon \eta_t - \hat{\sigma}_\varepsilon \sum_{j=1}^{q} \hat{\theta}_j \eta_{t-j}$ ----(3)

こうして，η_t を Monte Carlo 法により発生させれば，式(3)より新たな z_t が，さらに式(1)により y_t が算出されることになる。以下にこの模擬発生手順を示す。(a) η_1 を発生させ，式(3)の z_1 を決定する。ただし，z_0，$z_{-1}, \cdots, z_{-p+1}$ および $\eta_0, \eta_{-1}, \cdots, \eta_{-q+1}$ は，0とおく。この手順を $z_2, z_3, \cdots, z_{N'}$ が発生されまでくり返す。ここで，$N' = N_g + N_w$ であり，N_w は，初期値の z_t，η_t を0とした影響を取り除くための予備発生年数で，実際に必要とする発生年数が N_g である。　(b) 発生された $z_1, z_2, \cdots, z_{N_w}, \cdots, z_{N'}$ から最初の z_{N_w} を除き，残りの N_g 個である $z_{N_w+1}, z_{N_w+2}, \cdots, z_{N_w+N_g}$ を新たに $z_1, z_2, \cdots, z_{N_g}$ とする。　(c) $z_1, z_2, \cdots, N_g$ を式(1)に代入し，模擬年水文量 $y_1, y_2, \cdots, y_{N_g}$ を算出する。

(2) 予測手法　時刻 t の時点から L 時刻先の予測値 z_{t+L}，$L \geqq 1$ は，最小平均二乗誤差予測 $z_t(L)$ により推定できよう。　Box ＆ Jenkins (1976)[2] は，この $z_t(L)$ を，z_{t+L} の条件付期待値として次式の予測関数を定義している。　$z_t(L) = E[z_{t+L} \mid z_t, z_{t-1}, \cdots]$ ----(4)　これより，式(2)の予測関数は次式となる。

$z_t(L) = E[z_{t+L}] = \hat{\phi}_1 E[z_{t+L-1}] + \cdots + \hat{\phi}_p E[z_{t+L-p}] - \hat{\theta}_1 E[\varepsilon_{t+L-1}] - \cdots - \hat{\theta}_q E[\varepsilon_{t+L-q}]$ ----(5)

上式において，時刻 t とそれ以前の値は，すでに生起した値であり，それらの条件付期待値はそれら自身の値となる。すなわち $E[z_{t-j}] = z_{t-j}$，$E[\varepsilon_{t-j}] = \varepsilon_{t-j}$，$j = 0, 1, 2, \cdots$ となる。他方，時刻 $t+1$ 以後の将来値 z_{t+j}，ε_{t+j} の期待値は，$E[z_{t+j}] = z_t(j)$，$E[\varepsilon_{t+j}] = 0$，$j = 1, 2, \cdots$ となる。たとえば，1ステップ先 ($L=1$) の予測値は，次式となる。　$z_t(1) = \hat{\phi}_1 z_t + \cdots + \hat{\phi}_p z_{t-p-1} - \hat{\theta}_1 \varepsilon_t - \cdots - \hat{\theta}_q \varepsilon_{t-q+1}$ ----(6)

ここで，時刻 $t+L$ での観測値 z_{t+L} は，$t+L$ 以前のランダム変量 ε_t の重み付無限和で表わすことができ，さらにこれは，時刻 L 以降の将来値と，過去から時刻 t までの値とに分けて表わすことができる。すなわち式(2)は，次式となる。　$z_{t+L} = \varepsilon_{t+L} + \hat{\psi}_1 \varepsilon_{t+L-1} + \hat{\psi}_2 \varepsilon_{t+L-2} + \cdots = \sum_{j=0}^{L-1} \hat{\psi}_j \varepsilon_{t+L-j} + \sum_{j=L}^{\infty} \varepsilon_{t+L-j}$，$\psi_0 = 1$ ----(7)

こうして式(7)は，L 時刻先の z_{t+L} の最小平均二乗誤差予測 $z_t(L)$ および予測誤差 $e_t(L)$ から成っていることがわかる。すなわち　$z_{t+L} = e_t(L) + z_t(L)$　あるいは，$e_t(L) = z_{t+L} - z_t(L)$ ----(8) となっている。そこで，式(5)による予測値の精度は，式(8)の予測誤差 $e_t(L)$ の分散で表わされることになり，次式となる。

$Var[e_t(L)] = E[e_t^2(L)] = \hat{\sigma}_\varepsilon^2 (1 + \sum_{j=1}^{L-1} \hat{\psi}_j^2) = \sum_{j=0}^{L-1} \hat{\psi}_j^2 \hat{\sigma}_\varepsilon^2$，$\hat{\psi}_0 = 1$ ----(9)

また，実現値 z_{t+L} の予測値 $z_t(L)$ に対する $(1-\alpha)$ 信頼区間は，次式で表わすことができる。

$\left[z_t(L) - u_{\frac{\alpha}{2}} \sum_{j=0}^{L-1} \hat{\psi}_j \hat{\sigma}_\varepsilon^2 < z_{t+L} < z_t(L) + u_{\frac{\alpha}{2}} \sum_{j=0}^{L-1} \hat{\psi}_j^2 \hat{\sigma}_\varepsilon^2 \right]$ ------(10)　ここで，$u_{\frac{\alpha}{2}}$ は標準正規分布表から求まる値であり，$\hat{\psi}_j$ は，式(7)の $L=0$ とした式を，式(2)に代入して得られる $\sum_{j=0}^{p} \sum_{i=0}^{\infty} \hat{\phi}_j \hat{\psi}_i \varepsilon_{t-j-i} = \sum_{j=0}^{q} \hat{\theta}_j \varepsilon_{t-j}$, $\hat{\phi}_0 = \hat{\theta}_0 = 1$ ····(11) において，右辺と左辺は同値の ε_k ($k = t, t-1, t-2, \cdots$) に対し等しいとして算出される。

3. 適用例

岩手県胆沢川の石淵ダム（集水面積 154.4 Km^2）への年流入量データ（昭29 − 51，23年間）を用いて，図−1の手順に従い確立したモデルによるシミュレーションについては，以下のごとくである。

(1) データ　石淵ダムへの年流入量データは，表−1のごとくである。

表−1　石淵ダムへの年流入量（m^3/s）（昭29 − 55）

昭和	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
流量	3904.7	4721.6	4837.1	4564.86	4885.23	5135.53	4309.46	5173.03	4717.57	3970.06	4788.11	5836.32	5336.27	4231.21	4677.42	4451.20
昭和	45	46	47	48	49	50	51	52	53	54	55					
流量	4217.98	4207.26	4952.51	3868.21	6120.22	4108.91	4103.27	4502.96	3556.65	5218.38	4905.48					

(2) モデル式　表−1における昭和51年までの年流入量データを用い，図−1の手順によりモデル化を試みた結果，次式のARMA(0, 1)すなわちMA(1)モデルが確立できた。

$y_t = 4657.29 + 591.64\, z_t$ -----(12)　　$z_t = \varepsilon_t - 0.192\, \varepsilon_{t-1}$ ----(13)

(3) 模擬発生　式(13)の ε_t は，平均0，分散 $\hat{\sigma}_\varepsilon^2 = 0.946$ をもつ独立正規ランダム変量となっているので，$\zeta_t = \varepsilon_t / \hat{\sigma}_\varepsilon$ なる標準正規ランダム変量を導入し，この ζ_t を次式の Box & Muller (1958) 法[3] によって460個発生させた。　$\zeta_1 = \sqrt{2\ln(1/u_1)} \cdot \cos(2\pi u_2)$，　$\zeta_2 = \sqrt{2\ln(1/u_1)} \cdot \sin(2\pi u_2)$ ····(14)　ここで，u_1，u_2 は，平均0，分散1の標準一様乱数である。　こうして発生された460年間の模擬年流入量 y_t の基本統計量および遅れ時間 $k = 1, 2$ 年の自己相関係数 r_1，r_2 と，表−1の年流入量データのそれらとを比較して表−2に示す。

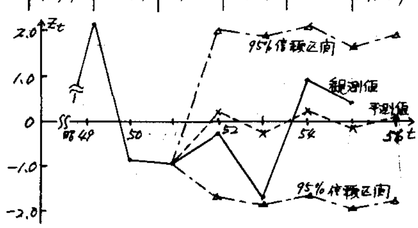
表−2　基本統計量および自己相関係数

データ＼統計量	平均 $\hat{\mu}$	標準偏差 $\hat{\sigma}$	歪み係数 $\hat{\gamma}$	r_1	r_2
年流入量	4657.29	591.64	0.840	−0.193	−0.143
模擬発生	4679.23	586.45	0.137	−0.144	−0.011

(4) 将来予測　式(13)の予測関数は，$z_t(L) = -0.192\, E[\varepsilon_{t+L-1}]$ ---(14) となる。この式より，$L=1$ においては $z_t(1) = -0.192\, \varepsilon_t$ となるが，$L \geqq 2$ では，$z_t(L) = 0$ となり予測値 y_t は平均値 $\hat{\mu}$ に収束することになる。こうしてARMA(0, 1)モデルでは，1ステップ先の予測のみが可能である。そこで昭和51年を現時点として昭和55年までの4年間の1年先の予測をし，その結果と観測値および有意水準 $\alpha = 0.95$ ($u_{\frac{\alpha}{2}} = 1.96$) に対する $\pm u_{\frac{\alpha}{2}} \hat{\sigma}_\varepsilon^2$ の算出結果を表−3に，またそれらを図−2に示す。

表−3　予測値，観測値および $\pm u_{\frac{\alpha}{2}} \hat{\sigma}_\varepsilon^2$

昭和	$\hat{\theta}_1$	ε_t	$\hat{\sigma}_\varepsilon^2$	予測値 $z_t(1)$	観測値 z_{t+1}	$\pm u_{\frac{\alpha}{2}} \hat{\sigma}_\varepsilon^2$
51	0.192	−1.032	0.946		−0.936	
52	0.174	−0.434	0.956	0.198	−0.255	±1.854
53	0.131	−1.785	0.977	0.076	−1.728	±1.874
54	0.226	0.564	0.928	0.234	0.967	±1.915
55	0.177	0.541	0.957	−0.127	0.442	±1.819



図−2　観測値，予測値および信頼区間

(5) 考察　確立されたモデル式(12)，(13)により発生された模擬データの統計的諸特性は，表−2からわかるように歪み係数の再現性には難があるものの，その他の量は概ね良好である。また将来予測においては，ここでのように次数 p，q が低次であり，しかも水文量の自己相関性が小さな場合には，L 時刻先の予測が不可能となる。また図−2のように1ステップ先の予測においても良好とは言い難い。

4. おわりに

ARMAモデルは，その理論および統計的処理方法がかなり簡明であり，またそのシミュレーションも合理的であるから，水資源計画に大いに利用しうると思われる。ただし適用例でわかるように，水文量がほぼ独立過程と見なせるような場合のシミュレーションにおいては，模擬発生手法はともあれ，将来予測手法そのものの検討が必要であろう。今後このことと，月単位の水文量のモデル化とそのシミュレーションを検討したい。

参考文献　1) 山本：第37回年講第II部 PP601-602，　2) Box & Jenkins, "Time Series Analysis" H-Day, 1976，3) 三浦："シミュレーション入門" オーム社