ARMA 多季節模型에 의한 河川流量의 發生

漢陽大 大学院 金時永
高麗大 教授 尹龍男

序 論

河川流量의 歷史的인 資料는 너무 짧기 때문에 將來에 대한 가뭄이나 홍수에 대한 모든 可能한 樣相을 포함시킬 수 없다. 推計学的 模型 (Stochastic model)에 의한 河川合成流量의 發生은 將來에 대한 水資源系統의 計劃이나 運用에 있어서 단지 歷史的인 資料만을 사용하는것 보다 좀더 포괄적으로 解析할 수 있는 時系列을 제공한다.

이러한 模型의 時系列形態 (type of time series)는 河川流量의 發生에 대하여 여러종류의 模型이 있어나 크게 短期模型 (short momery models)과 長期模型 (Long momery models) 두 種類의 模型으로 구분할 수 있다.

AR, ARMA 와 ARIMA 模型들은 短期模型에 속하는 반면에 Fractional Gaussian noise (FGN) 模型과 Broken line (B.L) Processes 模型들은 長期模型에 속한다.

本研究에서는 短期模型중 AR 模型 (年, 季節模型)과 ARMA (年, 季節模型)에 대하여 다루었다.

本研究의 目的은 이러한 模型들중 ARMA 年模型은 旣開發 되어 보편화 되어 있으나 ARMA 季節模型은 아직 初期段階에 있어 重点을 두었으며 이러한 模型을 河川流量의 發生에 적용시켜 歷史的인 資料와 發生된 資料로 부터 얻어진 統計的인 特性과 Correlogram에 대하여 比較檢討 한후 將來에 대한 河川流量의 發生에 적합한 模型을 選択하는데 基礎를 마련함에 있다.

1. Annual Model

(1) AR 模型

1) 基本方程式

K次의 Autoregressive(AR) 혹은 Markov 模型은 다음과 같이 표현된다

$$X_t = \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_k X_{t-k} + \varepsilon_t \qquad \text{——— (1)}$$

여기서 X_t : 自己回歸過程 (Autoregressive process)

$\beta_1, \beta_2 \cdots \beta_k$; 自己回歸係數 (Autoregressive coefficient)

ε_t ; 平均이 0 이고 分散이 σ_ε^2인 獨立時系列

(1)식에서 Logs AR(1) 模型에 대하여서는 Normal AR(1) 模型과 같이 표현된다.

$$X_t = \beta_1 X_{t-1} + \varepsilon_t \qquad \text{——— (2)}$$

여기서 $\varepsilon_t = z_t \sigma_\varepsilon$ 이고 z_t는 平均이 0 이고 分散이 1 인 均等分布된 亂數 (Normally distributed random number)

X_t의 標準化는 다음과 같이 표현된다.

$$X_t = \frac{Q_t - \bar{Q}_t}{S_{Q_t}} \qquad \text{——— (3)}$$

여기서 Q_t는 자연Log를 취한 資料의 流量, $\bar{Q}_t$는 平均 S_{Q_t}는 標準偏差이다.

2) 模型의 媒介変数 推定

平均, 標準偏差, Random成分의 分散은 資料로 부터 다음과 같이 구할 수 있다.

$$\bar{Q}_t = \frac{1}{N}\sum_{i=1}^{N} Q_t \qquad (4)$$

$$S_{Q_t}^2 = \frac{1}{N-1}\sum_{i=1}^{N} (Q_t - \bar{Q}_t)^2 \qquad (5)$$

$$\sigma_\varepsilon^2 = S_{Q_t}^2 (1-\rho_1^2) \qquad (6)$$

여기서 ρ_1는 資料로 부터 구한 Lag-one (r_1) 時系列 相関係数로부터 推定되며 다음과 같이 표현된다.

$$r_k = \frac{\frac{1}{N-1}\sum_{t=1}^{N} (Q_t - \bar{Q}_t)(Q_{t+k} - \bar{Q}_t)}{\frac{1}{N}\sum_{t=1}^{N} (Q_t - \bar{Q}_t)^2} \qquad (7)$$

(7)식에서 $k=1$ 일때 r_1을 구할 수 있다.

$$\beta_1 = \rho_1 \qquad (8)$$

3) 発生方程式 (Generating equation)

(2),(6),(7) 식을 이용하여 (3) 식을 역변환 하면

$$\begin{aligned} Q_t &= \bar{Q}_t + \beta_1 X_{t-1} S_{Q_t} + Z_t \sigma_\varepsilon \\ &= \bar{Q}_t + \rho_1 X_{t-1} S_{Q_t} + Z_t S_{Q_t}\sqrt{1-\rho_1^2} \qquad (9) \end{aligned}$$

$$Q_t^* = EXP(Q_t) \qquad (10)$$

(9)식에서 $\rho_1 = r_1$로 대치되고 초기값 $X_0 = 0$로 놓고 過程을 반복함으로서 계속적인 流量系列이 發生된다.

(2) ARMA 模型

1) 基本方程式

Mixed Autoregressive Moving Average Model, ARMA(p, q) 模型은 다음과 같이 표현된다.

$$q_t = \beta_1 q_{t-1} + \beta_2 q_{t-2} + \cdots + \beta_p q_{t-p} + \varepsilon_t + \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-2} + \cdots + \alpha_q \varepsilon_{t-q} \qquad (11)$$

여기서 $\beta_1, \beta_2 \cdots \beta_p$: 自己回歸係數

$\alpha_1, \alpha_2 \cdots \alpha_q$: MA 係數

ε_t; 平均이 0 이고 分散이 σ_ε^2인 獨立時系列

(11)식에서 ARMA(1, 1) 模型에 대하여 다음과 같이 표현된다.

$$q_t = \beta_1 q_{t-1} + \varepsilon_t + \alpha_1 \varepsilon_{t-1} \qquad (12)$$

(12)식에서 推定되어야 할 媒介變數는 $p+q+2$ 즉 4개 μ, α_1, β_1, Random 成分의 分散 σ_ε^2 이다.

2) 模型의 媒介變數의 推定

β_1과 α_1은 Yule Walker 方程式, 최소자승법 과 Maximum likely 方法에 의하여 推定되며 Box and Jenkins[3] 에 잘 설명되어 있으며 다음식은 Yule walker 方程式 이다.

$$\hat{\beta}_1 = \rho_2 / \rho_1 \qquad \text{——— (13)}$$

$$\hat{\rho}_1 = \frac{(\hat{\beta}_1 - \hat{\alpha}_1)(1 - \hat{\alpha}_1 \hat{\beta}_1)}{1 - 2\hat{\alpha}_1 \hat{\beta}_1 + \hat{\alpha}_1^2} \qquad \text{——— (14)}$$

$$\sigma_\varepsilon^2 = \frac{1 - \hat{\beta}_1^2}{1 - 2\hat{\alpha}_1 \hat{\beta}_1 + \hat{\alpha}_1^2} \qquad \text{——— (15)}$$

(14)식에서 $\hat{\alpha}_1$ 에 대하여 根을 구한후 stationary 條件인 $|\hat{\alpha}_1| \leq 1$ 에 대하여 解를 취한다[3]

3) 發生方程式

(10),(13) 과 (15)식을 이용하여 (12)식을 역변환 하면

$$Q_t = \bar{Q}_t + \hat{\beta}_1 S_{Q_t} (Q_{t-1} - \bar{Q}_t) + S_{Q_t} \sigma_\varepsilon Z_t - \hat{\alpha}_1 S_{Q_t} \sigma_\varepsilon Z_{t-1} \qquad \text{——— (16)}$$

$$Q_t^* = EXP(Q_t) \qquad \text{——— (17)}$$

2. Seasonal Model

(1) Thomas - Fiering 模型

1) 發生方程式

正常時系列(Stationarity)[1][8] 이라함은 時間에 따라 統計的 特性이 變化하지않는 時系列을 뜻한다. 時系列의 平均이 時間에 따라 變하지않는것을 1次 正常時系列, 平均과 分散이 變하지않는것을 2次 正常時系列 혹은 弱正常時系列(Week stationary) 平均, 分散과 共分散(covariance)이 變하지 않는 것을 엄격한 의미에서 强正常時系列(Strong - stationary)이라 한다. 一般的으로 水文量은 엄격한 의미에서 正常時系列을 만족할 수 없으므로 (18)식과 같은 方法으로 標準化 함이 바람직하다.

平均과 分散의 주기성분은 Young-Pisano[4] 方法에 의하여 $Q_{t,\tau}$로부터 제거될 수 있다.

$$Q'_{t,\tau} = \frac{Q_{t,\tau} - \bar{Q}_\tau}{S_{Q_{t,\tau}}} \qquad (18)$$

$Q'_{t,\tau}$는 平均과 分散에 있어서 正常時系列이고 $Q'_{t,\tau}$는 정확히 平均과 分散이 각각 0과 1 아니면 다음과 같이 變換이 필요하다.

$$g^*_{t,\tau} = \frac{Q'_{t,\tau} - \bar{Q}'_\tau}{S'_{Q_{t,\tau}}} \qquad (19)$$

여기서 $\bar{Q}_{t,\tau}$와 $S'_{Q_{t,\tau}}$는 τ 季節에 대한 $Q'_{t,\tau}$의 平均과 標準偏差이다. Rosner와 Yevjevich[5]는 $g^*_{t,\tau}$를 Standardized fitted 時系列이라 불렀다.

$$g^*_{i,\tau} = g^*_\tau + \frac{S^*_\tau r^*_\tau}{S^*_{\tau-1}}(g^*_{i,\tau-1} - g^*_{\tau-1}) + \varepsilon_{i,\tau} S^*_\tau \left(\sqrt{1-\rho^2_\tau}\right) \qquad (20)$$

여기서 $\bar{g}^*_{i,\tau}$: $g^*_{i,\tau}$의 τ 季節平均
r^*_τ : $\tau, \tau-1$ 季節사이의 相關係數.
S^*_τ : τ 季節의 標準偏差
$\varepsilon_{i,\tau}$: 均等分布를 갖는 乱數

(20)식으로부터 發生된 값을 (18)식 혹은 (19) 식에 의하여 역변환 한다.

$$Q^*_{t,\tau} = \bar{Q}'_\tau + g^*_{t,\tau} S'_{Q_{t,\tau}} \qquad (21)$$

$$Q^{**}_{t,\tau} = EXP(Q^*_{t,\tau}) \qquad (22)$$

Matlas[6]는 流量의 Moment를 유지하기 위한 方法을 제안하였다.

(2) ARMA(1.1) 模型.

1) 基本方程式

$$g^*_{t,\tau} = \beta_{1,\tau} g^*_{t,\tau-1} + \varepsilon_{t,\tau} - \alpha_{1,\tau}\varepsilon_{t,\tau-1} \quad \text{———} (23)$$

여기서 $\tau = 1, 2 \cdots w$, $w = 4$

標準化는 (18),(19)식과 같은 方法으로 한다.

2) 模型의 媒介變數 推定

Lag-k ($r_{k,\tau}$) 季節相関係數는 시간간격 τ에 대하여 다음과 같이 표현된다.[1],[7]

$$r_{k,\tau} = \frac{\frac{1}{N}\sum_{t=1}^{N}(g^*_{t,\tau} - g^*_{\tau})(g^*_{t,\tau-k} - g^*_{\tau-k})}{S^*_{\tau}\, S_{\tau-1}} \quad \text{—} (24)$$

여기서 $r_{k,\tau}$는 $\rho_{k,\tau}$의 推定된 資料의 季節相関係數이고 $\tau-k<1$ 일때 t 대신 $t-1$, $g^*_{t,\tau-k}$ 대신 $g^*_{t,\tau-k+w}$ 로 대치하고 S^*_{τ}는 $g^*_{t,\tau}$의 標準偏差이다.

$\hat{\beta}_{1,\tau}$ 와 $\hat{\alpha}_{1,\tau}$, Random 成分 $\varepsilon_{t,\tau}$의 分散은 다음과 같이 推定된다.

$$\hat{\beta}_{1,\tau} = \frac{\rho_{2,\tau}}{\rho_{1,\tau-1}} \quad \text{———} (25)$$

$$\hat{\alpha}_{1,\tau} = \hat{\beta}_{1,\tau} + \frac{1-(\hat{\beta}_{1,\tau}\cdot\rho_{1,\tau})}{\hat{\beta}_{1,\tau} - \rho_{1,\tau}} - \frac{\hat{\beta}_{1,\tau+1} - \rho_{1,\tau+1}}{(\hat{\beta}_{1,\tau} - \rho_{1,\tau})\cdot\hat{\alpha}_{1,\tau+1}}$$

——— (27)

(25)식에서 $\hat{\alpha}_{1,t}$ 에 대하여 根을 구한후 두개의 根중 Stationary 條件인 $|\alpha_{1,1}| < 1$ 에 대하여 解로 채택한다.

$$\sigma^2_{t-1}(\varepsilon) = \frac{\hat{\beta}_{1,t} - \rho_{1,t}}{\hat{\alpha}_{1,t}} \quad \text{——— (26)}$$

여기서 $\rho_{1,t}, \rho_{2,t}$ 값 (24)식에 의하여 推定된 값이다.

3) 発生方程式

$$g^*_{t,t} = \bar{g}^*_{t,t} + (\beta_{1,t} g^*_{t,t-1} + \varepsilon_{t,t} - \alpha_{1,t}\varepsilon_{t,t-1}) S_{g^*_{t,t}}$$

$$= \bar{g}^*_{t,t} + (\beta_{1,t} g^*_{t,t-1} + Z_{t,t}\sigma_{t-1} - \alpha_{1,t} Z_{t,t-1}\sigma_{t-2}\varepsilon_{t,t-1}) S_{g^*_{t,t}} \quad \text{——— (28)}$$

(3) ARMA(2,1) 模型

1) 基本方程式

$$g^*_{t,t} = \beta_{1,t} g^*_{t,t-1} + \beta_{2,t} g_{t,t-2} + \varepsilon_{t,t} - \alpha_{1,t}\varepsilon_{t,t-1} \quad \text{——— (29)}$$

標準化는 (18),(19)식과 같다.

2) 模型의 媒介変数 推定

媒介変数 $\beta_{1,t}, \beta_{2,t}, \alpha_{1,t}, \sigma^2_\varepsilon(t)$는 다음과 같이 推定된다.

$$\hat{\beta}_{1,t} = \frac{\rho_{2,t}\cdot\rho_{1,t-2} - \rho_{3,t}}{\rho_{1,t-1}\cdot\rho_{1,t-2} - \rho_{2,t-1}} \quad \text{——— (30)}$$

$$\hat{\beta}_{2,t} = \rho_{3,t}\cdot\rho_{1,t-1} - \rho_{2,t}\cdot\rho_{2,t-1} \quad \text{——— (31)}$$

$$\sigma^2_{t-1}(\varepsilon) = \frac{\hat{\beta}_{1,t} + \hat{\beta}_{2,t}\cdot\rho_{1,t-1} - \rho_{1,t}}{\hat{\alpha}_{1,t}} \quad \text{——— (32)}$$

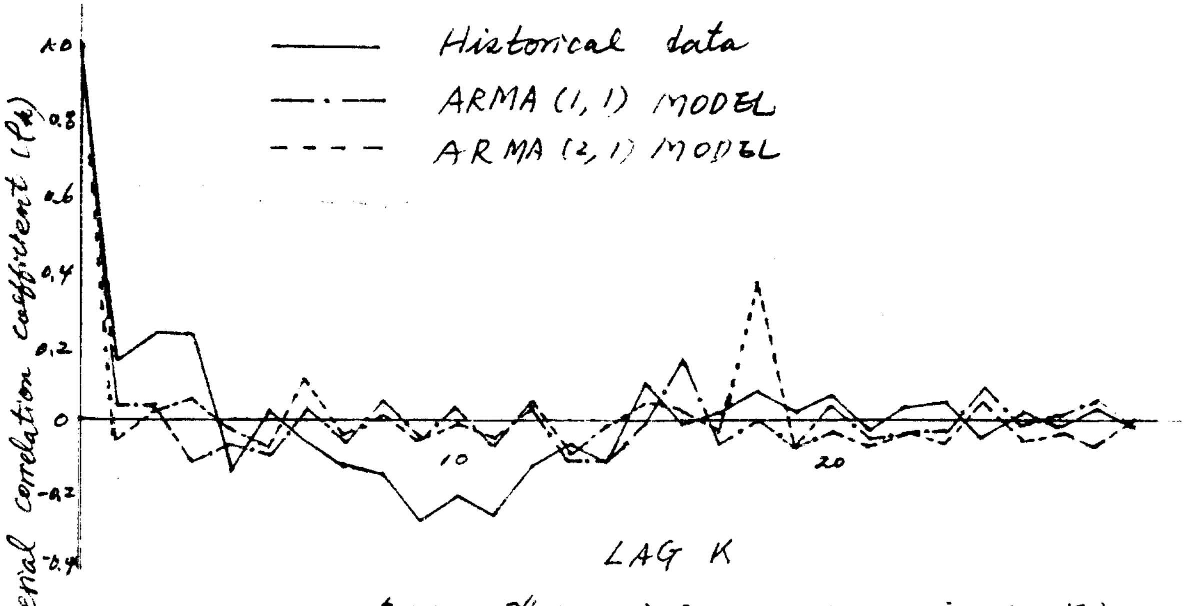

그림 3.1 資料와 發生된 流量의 Correlogram 比較 (年)

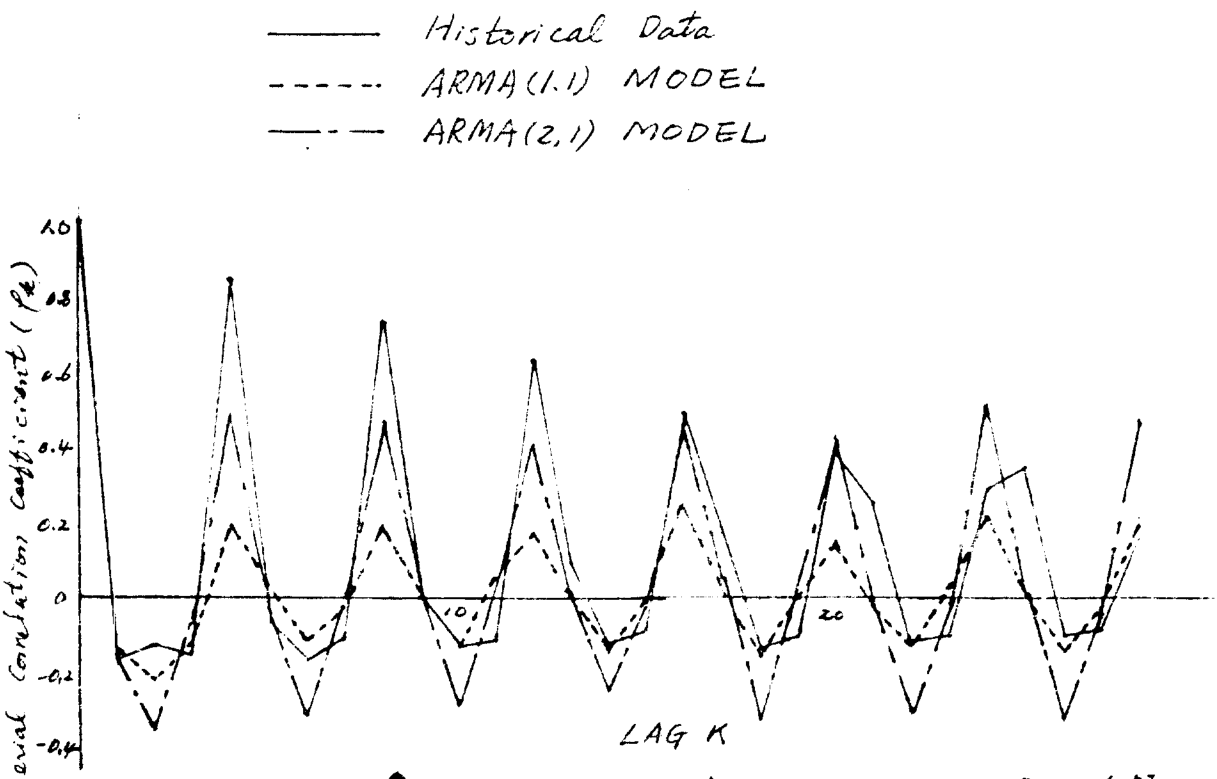

그림 3.2 資料와 發生된 流量의 Correlogram 比較 (季節)

3) 發生方程式

$$q^*_{t,\tau} = \bar{q}^*_\tau + (\hat{\beta}_{1,\tau} q_{t,\tau-1} + \hat{\beta}_{2,\tau} q^*_{t,\tau-2} + \xi_{t,\tau} \sigma_{\tau-1} - \hat{\alpha}_{1,\tau} \varepsilon_{t,\tau-1} \xi_{t,\tau-1}) S^*_{q^*_{t,\tau}} \qquad (33)$$

3. 模型結果分析

資料는 漢江流域의 淸平 1953 ~1982년 까지의 29年의 資料로서 各各의 模型에 대하여 200年間의 流量을 發生시켜서 發生된 流量의 初期 50年은 初期條件에 많은 영향을 받으므로 무시하고[2] 나머지 150年 發生된 流量에 대하여 Correlogram을 比較 하였다. 歷史的인 資料와 ARMA(1,1), ARMA(2,1) 模型의 年과 季節模型의 Correlogram은 그림(3.1)과 그림(3.2)에 各各 프롯트 되었다.

그림(3.1)은 季節流量으로부터 年流量으로 合成하여 구한 값들을 프롯트 한것이다.

후에 AR(年), Thomas-Fiering 과 ARMA(1,1) 模型들과 統計的 特性과 Correlogram을 比較함으로서 河川流量의 發生 模型의 選択에 좋은 資料가 될것이다.

참 고 문 헌.

1. J.D. Salas, J.W. Dellaur, V. Yejevich and W.L. Lane
"Applied Modeling of Hydrologic Time Series"
Water Resources Publications, 1980
2. Fiering, M.B and Jackson, B.B.
"Synthetic Streamflows"
Amer. Geophs. Union Water Res. Monograph 1, 1971
3. Box, G.E.P and Jenkins, G.M,
"Time Series Analysis Forecasting and Control", Holdenday, 1970
4. Young, G.K and W.C "Operational Hydrology using Residuals". J of Hyd. D.V. ASCE HY4 Jul., 1968
5. Roesener, L.A and Yevjevich V, M. "Mathematical Models for Time Series of Monthly Preciptation and Monthly Runoff" Hydrology paper No 15. Colorado state University, Fort.Collins, Colo, P50, 1966.
6. Matlas N, C "Mathematical Assessment of Synthetic Hydrology" Water Res. Res. P937-945, 1967.
7. Tao P.C and Deullem J.W "Seasonal and Nonseasonal ARMA Models in Hydrology" J of Hyd. Div. ASCE HY10 P1541~1559, 1976
8. N.T Kotteyoda "Stochastic Water Resource Technology", 1980.