

# 댐유입량 예측기법별 적정성 분석

## Analysis of Applicabilities for Dam Inflow Prediction Methods

○김주환\*, 강권수\*\*

### 1. 서 론

댐은 수자원의 효율적인 관리와 안정적인 용수공급을 위해 건설된 것으로서 댐유역의 수문특성에 따라 유출량의 변화가 다양하게 이루어진다. 특히 우리나라와 같이 홍수기와 이수기의 수문 특성 차이가 심한 곳에서는 홍수기에 저장된 물을 다음 홍수기까지 각종 용수로서 사용하게 되는데 안정적인 용수공급을 위해서는 댐유역으로 유입되는 예측유입량을 파악하여야 하며 이를 근간으로 각종 용수에 대한 공급계획이 수립될 수 있다. 그러나 미래의 사건을 예측하는 일은 기상 및 유역내 조건 등의 다양성으로 인한 불확실성을 포함하기 때문에 완벽한 예측을 수행하는 일은 언제나 완벽하지 않은 경우가 대부분이다. 이들 인자들은 시간간격, 상류 유역시스템의 특성, 유출을 발생시키는 강수의 특성이 있으며 정확한 유출량의 예측은 다음과 같은 인자들에 의존적이다. 즉 (1) 요구되는 선행시간에서 유역의 현행상태가 유량에 영향을 주는 정도, (2) 다른 입력(주로 강수)변수를 예측하는 정도이다. 그러나 예측을 수행하는 목적은 어느 정도의 오차를 수용하는 범위 내에서 오차를 최소화할 수 있도록 하는 것으로 수문 분야에서 가장 흔하게 이루어지고 있는 예측변수는 유출량과 강우량이다. 본 연구에서는 댐유입량의 예측에 적절한 방법은 어떠한 방법들이 있는지에 대하여 조사하고 이들 가운데 추계학적 모형과 시계열 분석모형, 상관계수법 그리고 최근에 예측기법으로 많이 적용되고 있는 신경망 모형에 대하여 소양강댐의 월유입량자료를 적용하여 검토하고자 한다. 소양강댐의 각종 월자료는 1974년부터 2000년도 자료를 이용하였으며 과거의 주요 패턴 및 시계열특성 등은 댐운영의 기초자료로 사용되고 있으며 이를 통하여 특히 이수기 각종 용수공급계획이 수립된다. 또한 월단위 댐유입량에 영향을 미칠 수 있는 기온, 습도, 증발산, 풍향, 풍속 등 기상자료 등을 조사하였으며 이들은 예측의 정확도를 얼마나 높일 수 있는지에 대한 검토를 위한 것이다.

### 2. 예측이론에 관한 고찰

수문특성 조사 및 모형화 단계에서 필요한 수문기상자료의 분석시 통계적인 방법의 문제점은 동적시스템으로 환경의 변화와 분석 대상자료에 포함된 잡음이며, 이로 인하여 주어진 문제에 대하여 고려해야 할 변수가 너무 많다는 것이다. 이들은 또한 정량적으로 표현하기 어려운 정성적 변수로서 작용하는 경우도 발생된다. 통계학적 분석방법을 통한 대표적인 모형으로서 중회귀분석 모형을 이용할 경우 변화에 대한 적용이 힘들고 분석자료가 갖는 분포에 대한 엄격한 가정이 요구되며 계량적 변수와 정성적 변수가 혼합된 경우 성과가 좋지 않은 것으로 알려져 있다.

\* 한국수자원공사 수자원연구소 선임연구원

\*\* 한국수자원공사 물관리센터 대리

시계열 분석을 통한 Box-Jenkins 모형은 주로 1년 이하의 단기예측에 이용되며, 장기예측의 경우에는 인과관계 분석을 통해 이루어진다. 시계열이란 일정한 시간간격을 두고 모아진 자료 즉, 시계열자료 그 자체에만 의존하여 예측을 행하는 것으로 시계열 생성과정을 블랙박스로 취급하고 단지 자료의 패턴변동에만 의존하여 이를 수학적 관계식으로 표현하고자 하는 것이다. 시계열 분석에서 자료의 패턴은 임의성(Randomness), 경향(Trend), 계절성(Seasonality), 주기(Cycle)로 나타나며 자료패턴을 파악하기 위해서는 일정한 지체시간(Time Lag)상에 위치하는 자료들간의 상호 의존관계를 나타내는 척도 즉, ACF(Auto-Correlation Function)를 이용한다.

근래 각종 자료의 예측이나 추정을 위해서 도입, 적용되고 있는 신경망은 자료가 가진 특성을 비선형 함수와 그 구조가 갖는 병렬적 특성을 이용해 시스템을 구성할 수 있다는 장점으로 인하여 널리 사용되고 있다. 신경망을 예측에 적용할 수 있는 것은 지금까지 사용되어온 어떠한 수학적 모형보다 구성이 용이하다는 점이다. 이는 일단 신경망이 예측에 필요한 정보를 가지고 있다면 어떠한 유사한 자료가 입력되더라도 학습시 적용된 가중치라는 매개변수를 통하여 출력값을 기억된 측정값들 가운데 내삽 등을 통하여 추정할 수 있기 때문이다.

### 3. 적용자료의 추계학적 분석

시계열 자료가 추계학적 분석은 각종 대상자료의 통계적 특성과 시계열 이론에서 제공되고 있는 방법에 의하여 모형화 할 수 있는데 특히 시계열 분석에 의한 Box-Jenkins 모형에서 AR, MA 차수 즉,  $p$ ,  $d$ ,  $q$ 를 결정하는 과정을 시스템 동정(System Identification)이라 하며 이를 위하여 다음과 같이 대상자료를 분석하였다. 이 결과는 각종 예측이론에 의한 모형의 구축에 변수를 선택하는 기준으로 작용한다.

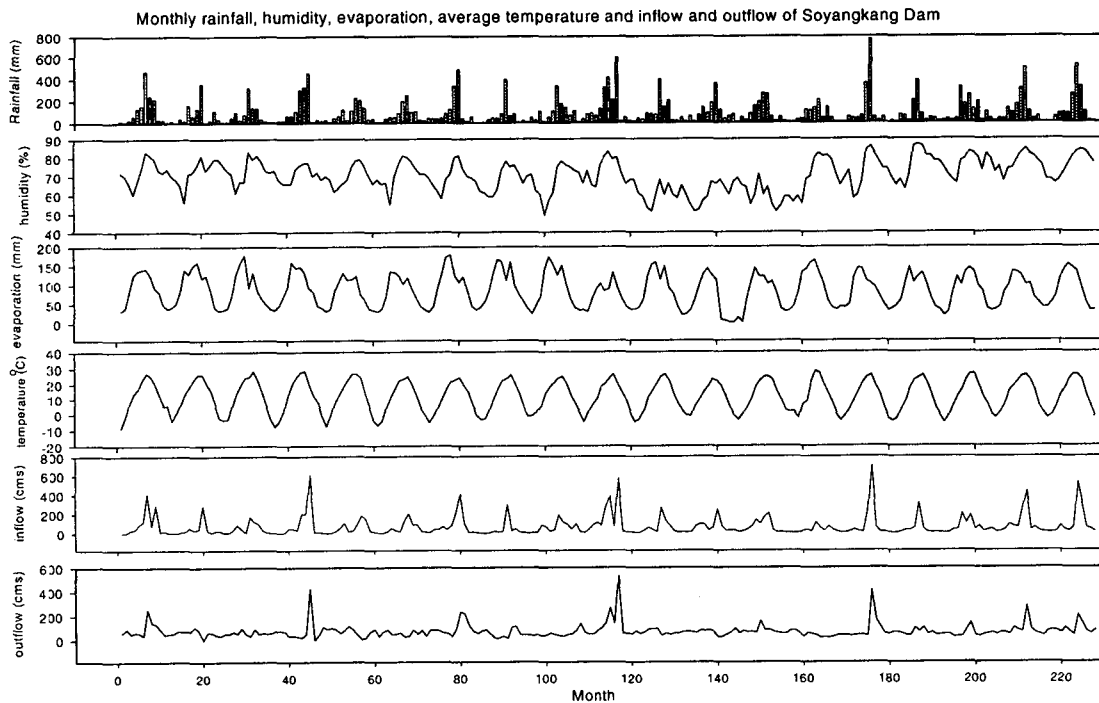


그림 1. 소양강댐 월강우량, 습도, 증발산, 평균기온, 월유입량 및 월방류량 계열 (N=324, 1974. 1 - 2000. 12)

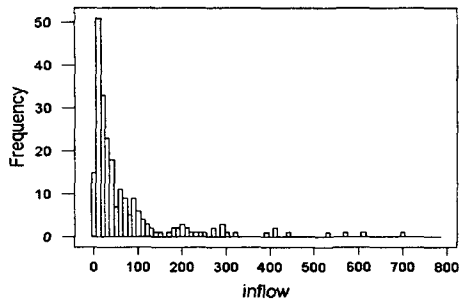


그림 2 월유입량 빈도분포

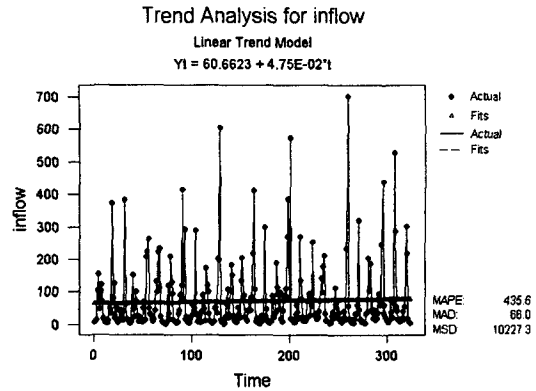


그림 3 월유입량의 경향

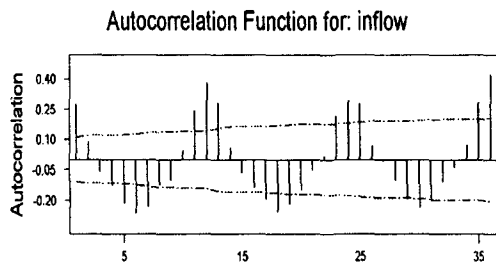


그림 4. 소양강댐 월유입량의 자기상관계수

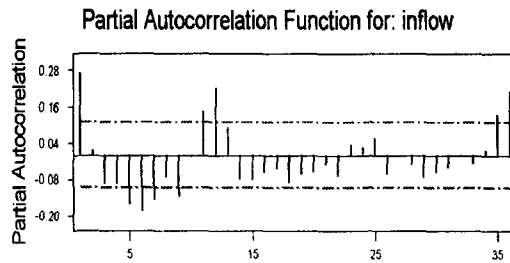


그림 5. 소양강댐 월유입량의 부분자기상관계수

그림 2와 3에서는 소양강 댐유입량에 대한 빈도 및 시간에 따른 경향을 나타낸 것으로 시간의 경과에 따라 약간씩 증가하는 경향을 보여 주었다. 또, 시계열 특성으로서 원 계열과 임의 시점만큼 떨어진 자료간의 상관성을 평가하는 자기상관함수의 변화양상을 도시하여 주기성을 그림 3에서와 같이 확인할 수 있다. 또한 자기상관계수를 이용하면 부분자기상관계수(Partial Auto Correlation Function; PACF) 라는 또 다른 여러 종류의 숫자를 계산할 수 있는데 이는 시계열 자료가 갖는 과거와 현재의 관계를 표현하는 또 다른 방법으로 추가적으로 새로운 정보를 주는 것이 아니고 자기상관계수를 조금 변형하여 새롭게 표현하는 것이다. 즉 부분 자기상관계수는 자기상관계수의 상관계수라 할 수 있다.

#### 4. 모형검토

##### 4.1 시계열 모형

시계열 모형에서는 10가지 승법 ARIMA(Multiplicative Auto-Regressive Integrated Moving Average)모형을 검토하여 잔차(Residual)의 SS(Sum of Square)와 MS(Mean Square)오차를 계산하여 오차가 최소인 시계열 모형을 선정하였으며 검토한 모형은 표 1에서 볼 수 있으며 여기서 선정된 모형의 매개변수는 표 2에서와 같이 추정되었으며 이를 원계열과 비교한 그림은 그림 6과 같다.

표 1 시계열 모형 검토결과

Model	SSR(Sum of Square of Residual)	MSR (Mean Square of Residual)	DF (Degree of Freedom)	Remark
ARMA(1,0,1)×(1,1,2)	1960415	6407	306	
ARMA(1,0,1)×(2,1,2)	<b>1871382</b>	<b>6136</b>	305	○
ARMA(1,0,1)×(2,2,2)	2259847	7713	293	
ARMA(1,0,1)×(1,2,1)	3229743	10948	295	
ARMA(1,1,0)×(1,1,1)	2942438	9584	307	
ARMA(1,0,1)×(2,1,1)	1902770	6218	306	
ARMA(1,0,1)×(2,1,0)	2312130	7531	307	
ARMA(2,0,0)×(2,1,1)	1902022	6216	306	
ARMA(2,0,1)×(2,1,0)	2311329	7553	306	
ARMA(2,0,1)×(2,1,1)	1901633	6235	305	

표 2. 소양강댐 유입량 예측모형 ARMA(1,0,1)×(2,1,2)에 대한 매개변수 추정

구분	Parameters	STD	t-ratio	Remark
AR 1	-0.1936	0.9042	-0.21	
SAR 12	-0.5168	0.1961	-2.64	
SAR 24	-0.2545	0.0622	-4.09	
MA 1	-0.1329	0.9144	-0.15	
SMA 12	0.4905	0.1996	2.46	
SMA 24	0.4130	0.1983	2.08	
Constant	1.1512	0.6675	1.72	

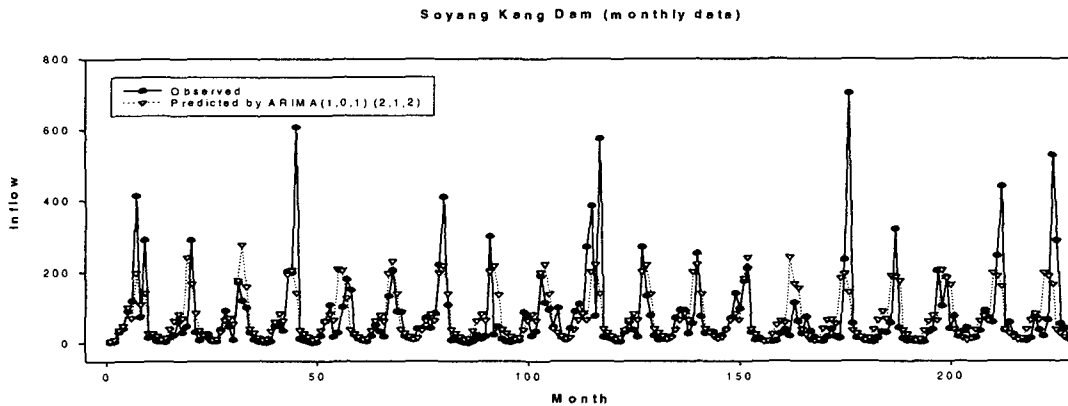


그림 6. 시계열모형 ARMA(1,0,1)×(2,1,2)에 의한 결과 비교

#### 4.2 중회귀 모형

중회귀 모형을 이용한 댐유입량 예측에는 유입량, 강우량 및 증발산량자료를 사용하였으며 3가지 경우에 대하여 추정된 중회귀식 모형식과 함께 다음과 같이 살펴보았다.

(1) 모형 1 (Multi\_R\_1)

모형구조  $I_{t+1} = F(I_t, I_{t-1}, I_{t-2}, R_t, R_{t-1}, R_{t-2}, E_t, E_{t-1}, E_{t-2})$

$$\begin{aligned} \text{추정식 } I(t+1) = & - 5.5 - 1.04 I(t) - 0.104 I(t-1) + 0.179 I(t-2) + 1.17 R(t) \\ & + 0.163 R(t-1) - 0.195 R(t-2) + 0.603 E(t) - 0.028 E(t-1) - 0.209 E(t-2) \end{aligned}$$

(2) 모형 2 (Multi\_R\_2)

모형구조  $I_{t+1} = F(I_t, I_{t-1}, I_{t-2}, R_t, R_{t-1}, R_{t-2}, E_t)$

$$\begin{aligned} \text{추정식 } I(t+1) = & - 12.2 - 0.972 I(t) - 0.036 I(t-1) + 0.192 I(t-2) + 1.09 R(t) \\ & + 0.082 R(t-1) - 0.231 R(t-2) + 0.545 E(t) \end{aligned}$$

(3) 모형 3 (Multi\_R\_3)

모형구조  $I_{t+1} = F(I_t, I_{t-1}, I_{t-2}, R_t, R_{t-1}, R_{t-2})$

$$\begin{aligned} \text{추정식 } I(t+1) = & 25.3 - 1.10 I(t) - 0.033 I(t-1) + 0.231 I(t-2) + 1.26 R(t) \\ & + 0.131 R(t-1) - 0.291 R(t-2) \end{aligned}$$

여기서, I는 유입량, R은 강우량, E는 증발산량을 나타내며 그 결과는 그림 7에 도시하였다..

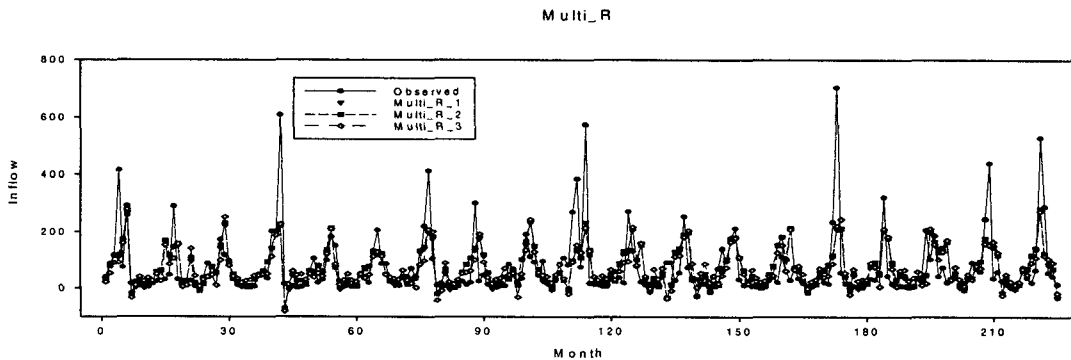


그림 7. 중회귀 모형에 의한 결과 비교

### 4.3 신경망 모형

신경망 모형은 시계열 모형과 중회귀 모형에서 유입량 예측에 필요한 변수를 선정하였듯이 이와 유사한 맥락을 갖는다. 단, 자료들이 포함하고 있는 잡음과 각 변수가 가지는 크기를 정규화하기 위한 전처리 과정과 계산 후의 후처리 과정이 신경망 모형을 적용하기 전후에 포함되며 예측에 적합한 모형을 선택하는 일은 아직까지 시행착오에 의한 반복적인 과정이 요구된다. 그러나 어떠한 구조를 갖는 신경망 모형일지라도 그 모형구조에 적합한 가중치를 찾아냄으로써 모형구조의 선택이 자유로운 장점을 가지고 있다. 신경망을 이용한 댐유입량의 예측은 다음 구조의 3가지 모형에 대하여 검토하였으며 본 연구에서 고려한 각 수문기상변수들을 조사하여 사용가능한 자료를 선택, 이들의 시계열 특성을 분석하여 입력변수를 설정하였다. 또한 3가지 모형에 의한 예측 결과는 그림 8에 도시하였다.

(1) 모형 1 (Model I)

$$I_{t+1} = ANN(I_t, I_{t-1}, I_{t-2}, R_t, R_{t-1}, R_{t-2})$$

(2) 모형 2 (Model II)

$$I_{t+1} = ANN(I_t, I_{t-1}, I_{t-2}, R_t, R_{t-1}, R_{t-2}, E_t, E_{t-1}, E_{t-2})$$

(3) 모형 3 (Model III)

$$I_{t+1} = ANN(I_t, I_{t-1}, I_{t-2}, I_{t-3}, I_{t-4}, I_{t-5})$$

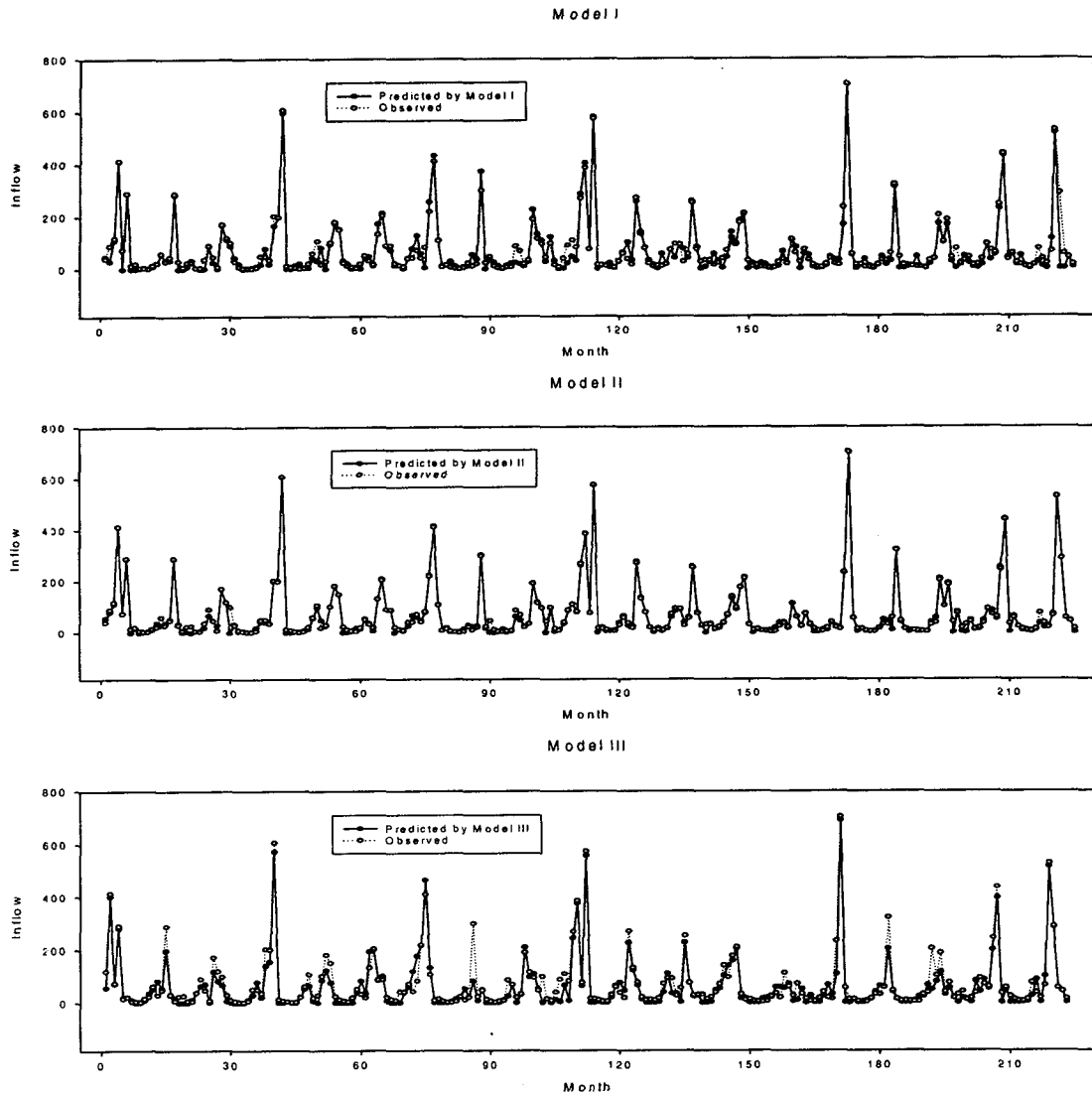


그림 8. 신경망 모형에 의한 예측결과 비교

## 5. 결 론

본 연구는 다목적댐 운영에 있어서 갈수에 대한 수문자료의 특성을 분석하여 물관리의 효율화 및 예측적 댐 운영을 제고시키기 위하여 소양강 유역을 대상으로 유입량 예측에 관한 연구내용을 기술한 것으로 결정론적 방법과는 대조적으로 관련된 내적인 물리적 과정을 고려하지 않고 현상 간의 관계를 확립하기 위해 지금까지 사용되어 온 추계학적 방법으로서 통계적 분석을 기초로 하는 중회귀모형, 시계열 모형 및 신경망모형에 대하여 고찰하였다. 예측은 일반적으로 과거의 경험 및 자료를 기반으로 한다. 본 연구에서는 최근 많이 적용되고 있는 신경망모형의 예측성능이 가장 양호한 결과를 보여주고 있으며, 또한 본 연구의 방법론을 타 수계에 적용하는 것은 비교적 용이하게 이루어 질 수 있다. 본 연구를 통하여 검토된 결과가 타 유역에 대해서도 그대로 재현된다고 단정적으로 말할 수는 없으나 지금까지의 방법과는 큰 차이가 발생하지 않을 것으로 생각된다. 다만 타 수계에 확장 적용을 시도하고자 할 때에는 적절한 검증 작업과 예측시스템의 구축이 이루어져 현재까지의 방법 중에서 최적 예측이 이루어질 수 있도록 하여야 할 것이다.