

용존산소 농도모의시 상태공간모형과 승법ARIMA모형의 시계열 분석

이원호 · 서인석* · 한양수**

충주대학교 건설도시공학과 · 수자원공사 상하수도 연구부* · 대구대학교 토목공학과**

Time series Analysis of State-space Model and Multiplication ARIMA Model in Dissolved Oxygen Simulation

Woon-Ho Lee · In-Seok Seo* · Yang-Su Han**

Dept. of Construction & Urban Engineering, ChungJu National University

* *Water Supply & Sewerage Team, Water Resources Research Institute, KOWACO*

** *Dept. of Civil Engineering Taegu University*

Abstract

The purpose of this study is to develop the stochastic stream water quality model for the intake station of Chung-Ju city waterworks in the Han river system.

This model was based on the theory of Box-Jenkins Multiplicative ARIMA(SARIMA) and the state space model to simulate changes of water qualities. Variable of water qualities included in the models are temperature and dissolved oxygen(DO). The models development were based on the data obtained from Jan. 1990 to Dec. 1997 and followed the typical procedures of the Box-Jenkins method including identification and estimation. The seasonality of DO and temperature data to formulate for the SARIMA model are conspicuous and the period of revolution was twelve months.

Both models had seasonality of twelve months and were formulates as SARIMA $(2,1,1)(1,1,1)_{12}$ for DO and temperature. The models were validated by testing normality and independency of the residuals.

The prediction ability of SARIMA model and state space model were tested using the data collected from Jan. 1998 to Oct. 1999. There were good agreements between the model predictions and the field measurements.

The performance of the SARIMA model and state space model were examined through comparisons between the historical and generated monthly dissolved oxygen series. The result reveal that the state space model lead to the improved accuracy.

I. 서 론

환경기초시설의 꾸준한 건설과 환경보전에 대한

인식의 확산으로 4대강 본류하천의 수질은 악화되고는 있으나 그 변동폭이 크지 않다. 그 동안 각종 환경 및 수자원 계획시 짧은 기간동안 수질 개선

효과가 큰 하수처리장, 분뇨처리장 등 계획을 위하여 단기간의 수질예측 기법인 확정론적 방법이 주로 사용되어 왔다. 물론 장기간의 추계학적 기법을 활용하기 위해서는 장기간의 측정자료가 소요되므로 통계학적 장기예측 방법이 실무에 적용하기 어려운 면이 없지 않았다. 이제 자료기간도 어느 정도 확보되었고 인위적인 환경질의 변화폭도 어느 정도 안정되었으며 환경관리 전문가시스템이나 의사결정 지원시스템 등에서는 확률론적 방법이 활발히 적용될 것으로 기대된다. 수질 및 수자원의 효율적 운영이나 계획, 설계 등을 위해서는 장기간의 월수질 및 수문자료에 대한 정확한 분석이 필요하다.

그러나 한정된 기간의 짧은 수질기록을 갖는 영역에서는 시스템의 거동예측이나 시스템 조작기준의 설정 등을 정확하게 할 수 없으므로 가능한 짧은 수질기록의 통계학적 특성을 유지시켜 장기간의 수질자료를 얻을 수 있는 추계학적 모의발생기법이 요구된다^{1~4)}.

본 연구의 방법은 충주시 상수원인 달천의 취수지점에서 유역을 하나의 폐쇄된 시스템으로 보고 월 용존산소를 시스템의 출력치, 그리고 월수온을 입력치로 간주하였으며, 본 연구에서 월 수질의 예측을 위해 선택된 수질자료는 취수지점에서 수년간 관측된 월평균 용존산소(DO)와 수온자료를 이용하였다. 이와 같이 선택된 월 수질자료는 승법 ARIMA모형과 상태공간 모형의 모수를 산정하는데 사용하였고 모형의 예측치와 실측값의 검증에 사용한다.

따라서 본 연구의 방향은 월 용존산소와 월 수온계열에 대한 경향적, 주기적 특성을 파악하고, 이에 적합한 단변량 시계열모형을 유도하고 월 수온 자료에 대해서도 동일한 방법의 모형을 분석하고, 상태공간모형을 유도하고자한다.

그리고 이와 같이 분석된 모형에 대하여 관측된 수온자료와 용존산소자료를 근거한 예측방법을 연구하고, 그 결과치를 단변량 시계열 모형인 승법 ARIMA모형과 다변량 시계열 모형인 상태공간모형의 결과치와 비교 검토하고자 한다.

II. 모형의 이론적 배경

단일변량 시계열에 대한 박스-젠킨스의 승법 ARIMA(p,d,q)(P,D,Q)₁₂ 모형은 모형의 식별, 추정, 검진의 3단계에 걸쳐 진행되는 시계열분석 및 예측기법이다. 시계열의 구성요소가 시간의 흐름에 따라 빠르게 변동할 때 유용한 모형이다.

상태공간모형은 주된 분석대상의 시계열데이터 Y_t 를 다른 시계열 데이터들을 사용해서 좀더 효과적으로 분석하기 위한 모형중 하나이다.

본 논문에서 승법 ARIMA(p,d,q)(P,D,Q)₁₂ 모형과 상태공간 모형에 대한 모형의 식별, 추정, 검진은 통계패키지 SAS/ETS를 사용하여 분석하였다⁵⁾.

1. 승법 ARIMA모형

승법 ARIMA모형은 1차, 2차 혹은 일반적으로 d차로 표현되는 단순 ARIMA(p,d,q)모형(비계절성 ARIMA모형)과 주기가 w인 계절성 차를 이용한 주기성 ARIMA(P,D,Q)모형(계절성 ARIMA모형)의 결합으로 표현될 수 있다.

주기가 w인 D차 계차조작(differencing)이 고려된 ARMA(P,Q)모형, 다시 말해서 ARIMA(P,D,Q)모형은 다음과 같이 된다.

$$(1 - \phi_1 B^w - \phi_2 B^{2w} - \dots - \phi_p B^{pw})(1 - B^w)^D Y_t = (1 - \theta_1 B^w - \theta_2 B^{2w} - \dots - \theta_q B^{qw}) a_t \quad (2.1)$$

여기서 B는 후향연산자(Backward Operator)로서 $BY_t = Y_{t-1}$, $B^w Y_t = Y_{t-w}$, $B^{pw} Y_t = Y_{t-pw}$ 등을 표시하며,

$$(1 - \phi_1 B^w - \phi_2 B^{2w} - \dots - \phi_p B^{pw}) = \Phi(B^w) \text{와} \\ (1 - \theta_1 B^w - \theta_2 B^{2w} - \dots - \theta_q B^{qw}) = \Theta(B^w) \text{는}$$

각각 seasonal AR 및 MA operator라고 부른다.

그리고 잔차 a_t 에 적용되는 ARIMA(p,d,q) 모형 즉, a_t 계열에 d차 계차조작(differencing)이 적용된 ARMA(p,q)모형은 마찬가지로

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - B)^d Y_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \quad (2.2)$$

여기서 등식 좌·우변의 괄호 속의 표현은 각각 $\phi(B)$ 및 $\theta(B)$ 로 표시되고 AR 및 MA 연산자라 하며 α_t 는 독립변수이다. 일반적인 다차항 ARIMA $(p,d,q) \times (P,D,Q)_w$ 모형은 식(2.2)의 α_t 를 식(2.1)에 대입하여 얻을 수 있으며 일반적으로 표시하면 다음과 같다⁵⁻⁷⁾.

$$\begin{aligned} & \Phi(B^w)\phi(B)(1-B^w)^D(1-B)^dX_t \\ & = \Theta(B^w)\theta(B)\alpha_t \end{aligned} \quad (2.3)$$

2. 상태공간모형

시스템의 상태(state)는 미래를 예측하기 위해서 필요한 현재와 과거정보(자료)의 최소집합으로 정의된다. 따라서 시스템의 미래현상을 현재상태와 미래의 입력을 사용해서 설명할 수 있다. 즉 어느 시스템의 상태벡터(state vector)는 미래의 값을 예측하기 위해서 필요한 현재와 과거의 정보(자료)의 최소량으로 정의된다. 그러므로 시스템의 미래 출력 벡터(out-put vector)는 현재의 상태벡터(state vector)와 미래의 입력벡터(input vector)를 사용하여 계산할 수 있다. 상태공간모형은 현재의 상태가 주어졌을 때 시스템의 미래는 현재상태에 의존한다는 Markov 성질을 바탕으로 하므로 마야코비안

표현이라 부르기도 한다.

시스템이 선형이고 시간에 불변(time invariant) 일 때, 이 시스템은 다음과 같은 상태방정식(state equation)과 출력방정식(out-put equation)으로 이루어진 상태공간모형으로 표현할 수 있다.

정상적 d변량 시계열 $\{y_t\}$ 가 VARMA 과정이면 다음과 같은 상태공간 모형으로 나타낼 수 있다.

$$\begin{aligned} Z_{t+1} &= Fz_t + GX_{t+1} \\ Y_t &= Hx_t + \xi_t \end{aligned} \quad (2.4)$$

단, z_t 는 $k \times 1$ 인 상태벡터(state vector)

F 는 $k \times k$ 인 변환행렬(transition matrix)

G 는 $k \times d$ 인 입력행렬(input matrix)

X_t 는 $n \times 1$ 인 시스템의 입력벡터(input vector)

y_t 는 $d \times 1$ 인 출력벡터(out-put vector)

H 는 $d \times k$ 인 출력행렬(out-put matrix) 또는 관찰행렬(observation matrix)

ξ_t 는 $d \times 1$ 인 관찰잡음(observation noise)

이 상태공간모형의 첫번째 방정식을 상태방정식(state equation) 또는 시스템방정식(system equation)이나 변환방정식(transition equation)이라 하

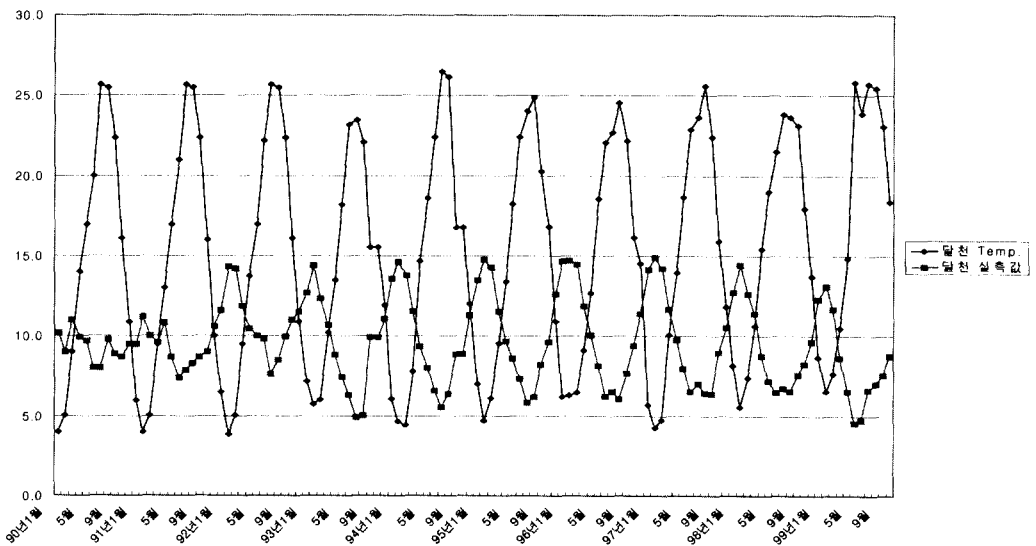


Fig. 1. Time serie plot of DO and water temperature in Dalchun station(1990~1999)

고, 두번째 방정식을 출력방정식 또는 측정방정식 (measurement equation)이나 관찰방정식(observation equation)이라 부른다.

III. 유역의 특성 및 적용자료

연구 대상하천인 달천은 한강 유역의 남부에 위치한 남한강의 제1지류 중의 하나로서 수계내 중·상류부인 동경 127° 10' ~127° 57' 북위 36° 29' ~36° 47' 사이에 위치하며 충북 보은군 내속리면과 경북 상주군 화북면의 도계인 속리산 천황봉에서 발원하여 청원군과 괴산군을 거쳐 내륙 도시인 충주시를 관통하여 칠금동에서 남한강 본류와 합류한다.

본 연구에서는 유역을 하나의 폐쇄된 시스템으로 보고 월 하천수질 중에서 월 용존산소(DO)와 월 수온 입력 값으로 간주하였다.

본 연구에서 월 수질의 예측을 위해 선택된 수질자료는 달천의 하류지점인 충주시 취수지점의 월 용존산소(DO)와 수온자료를 사용하였다.

월 수질자료의 취득은 일 평균 수질자료에 의한 월별로 집계하여 얻은 것을 사용하였으며 달천지점에서 10년간 관측된 월 평균 수질자료를 분석하여 사용하였다.

이와 같이 선택된 월 수질자료는 1990년부터 1997년 자료까지 12년간 모형의 모수를 산정하는

데 사용하였고 1998년 1월부터 1999년 10월까지의 모형의 예측치와 비교하였다. 그리고 Fig. 1.은 달천지점의 월별 용존산소(DO)와 월 수온이다.

IV. 모형의 적용 및 예측

1. 승법ARIMA 시계열모형

분석지점에서의 월 수질자료인 월 평균 용존산소(DO) 계열의 정상성을 판단하기 위하여 비계절성 계차(d=1)와 계절성 계차(D=1)를 고려하여 자기상관도(ACF)와 편자기상관도를 Fig. 2.에 나타내었다.

월 용존산소(DO) 계열은 정상시계열에 가까우며 그리고 지속성이 강하고 주기 w=12의 뚜렷한 계절성을 갖는 비교적 비정상성을 이루고 있다.

월 용존산소(DO) 계열의 계절성분과 경향성분을 제거하여 정상시계열로 바꾸고자 경향성분 및 계절성분에 대한 차분을 차례로 실시한 그림 4.1에서 ACF 및 PACF를 구하였다. 이로부터 월 용존산소(DO) 계열은 경향성분 및 계절성분의 차분차수를 각각 d=1 및 D=1로 할 때 정상시계열로 바꿀 수 있음을 알 수 있다.

월 용존산소(DO) 계열의 비계절(d=1)과 계절(D=1) 계차 변환된 ACF 및 PACF를 이용한 Box와 Jenkins의 모형판별방법으로부터 ACF유형이 시차 1과 계절주기 1, 12의 주기가 되는 시차에서

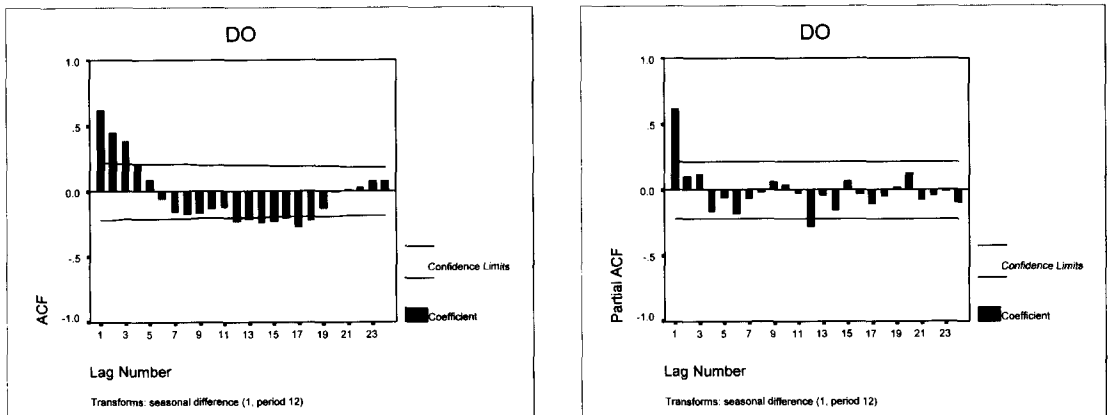


Fig. 2. Estimate ACF and PACF of DO by seasonal(D=1) and nonseasonal(d=1) differences in Dalchun station

0이 아닌 약간 큰 값을 갖는 불규칙적인 유형의 돌출점을 보이고 그 밖의 시차에서는 0인 값을 가지고 PACF유형을 보면 시차1, 12의 시차에서 0이 아닌 값을 갖는 불규칙적인 유형의 돌출점을 보이고 그 밖의 시차에서는 0인 값을 가지므로 승법 ARIMA(2,1,1)(1,1,1)₁₂모형으로 판별되었다.

월 용존산소(DO) 승법 ARIMA(2,1,1)(1,1,1)₁₂모형을 식으로 표현하면 다음과 같다.

$$(1+0.6909B+0.3135B^2)(1+0.12917B^{12}) \\ (1-B)(1-B^{12})Y_t \\ = (1+0.4082B)(1-0.0649B^{12})a_{yt} \quad (4.1)$$

월 용존산소(DO) 승법 ARIMA(2,1,1)(1,1,1)₁₂모형의 경우 $\sum a_{yt}^2(\phi, \theta, \theta)$ 면이 최소면을 이루게 되면 비선형 추적법에 의해 최종치인 비계절 AR(1), AR(2), MA(1) 및 계절SAR(1), SMA(1)의 모수를 구하였다.

잠정적으로 설정된 모형의 모수들을 추정하고 추정된 모형이 적절한지에 대한 검증 단계는 일반적으로 매개변수의 확인 검산을 위해 많이 이용하는 방법으로 설정된 모형에 의해서 추정된 잔차의 분석으로 모형이 잘 부합되었다면 잔차는 백색잡음의 성질을 만족시켜야 한다. 즉 모든 시점에 대하여 추정된 잔차의 평균이 0이고 분산이 일정해야

한다. 이를 알아보기 위해서 월 용존산소(DO) 승법 ARIMA(2,1,1)(1,1,1)₁₂모형으로 잔차를 추정하여 ACF와 PACF를 구하여 Fig. 3에 나타냈었다.

월 용존산소(DO) 승법 ARIMA(2,1,1)(1,1,1)₁₂모형에 의해 추정된 잔차의 결과 모든 시차에서 크게 유의한 값이 없으므로 추정된 잔차는 백색잡음의 성질을 만족한다고 볼 수 있다. 따라서 설정된 모형은 적절하다고 할 수 있으며 잠정 설정된 월 용존산소(DO) 승법 ARIMA(2,1,1)(1,1,1)₁₂모형은 장래 수질예측에 사용될 수 있다.

2. 상태공간모형

달천지점의 월 DO농도의 예측모형을 구축하기 위한 입력자료는 달천지점에서 DO농도에 영향을 주는 월 수온자료와 달천지점의 월 용존산소(DO)를 사용하였다. 상태공간 모형을 적용하기 전에 단변량 ARMA모형에서 사용한 방법과 같이 주어진 시계열 데이터가 정상적인 시계열 자료인지 확인해야 한다.

따라서 본 연구에서는 월 용존산소(DO)와 월 수온에 대한 비계절 차분 1과 계절 차분 1을 한 결과 정상성을 보이는 것으로 나타났다. 여기서 계절주기는 12월로 하였다.

달천지점의 DO예측을 위해 상태공간 모형을 구축을 위해 SAS/ETS프로그램을 수행한 결과 정준 상관분석을 통해 초기 VAR(p)모형은 Table 1.에

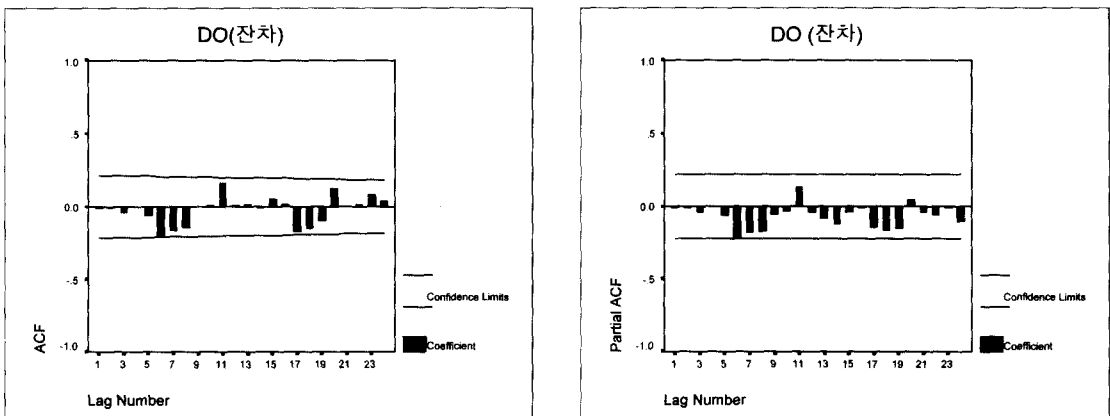


Fig. 3. ACF and PACF of residual estimated by multiplicative ARIMA(2,1,1)(1,1,1)₁₂ model(DO)

서의 아카이키계수(AIC)값이 가장 작은 값(102.54)을 갖는 시차는 시차 2으로 나타났기 때문에 VAR(2)모형으로 선정하였다.

Table 1. AIC values for select of P

Lag=0	Lag=1	Lag=2	Lag=3	Lag=4	Lag=5
115.076588	110.357325	102.539853	104.906962	109.376532	107.798514
Lag=6	Lag=7	Lag=8	Lag=9	Lag=10	
115.012517	115.552384	117.286567	123.192481	129.307007	

정준 상관계수를 이용하여 상태공간 z_t 의 구성요소를 찾아내는 절차를 보여주고 있다. $X1_t$, YT_t 를 각각 차분을 취한 t 시점에서의 달천지점의 월 평균 수온과 월 평균 용존산소(DO)농도라 할 때, 자료공간은 식(4.2)와 같다.

$$D_t = \{X1_t, YT_t, \dots, X1_{t-2}, YT_{t-2}\} \quad (4.2)$$

예측공간은 다음 식과 같다.

$$F_{t,1}^1 = \{X1_t, YT_t, \dots, X1_{t+2|t}, YT_{t+2|t}\} \quad (4.3)$$

Table 2.의 첫번째 행에서 잠정상태벡터 $\{X1_t, YT_t\}$ 에 새로운 원소 $\{X1_{t+1|t}\}$ 를 첨가 시켰을 때 최소정준 상관계수는 0.4291이던 DIC값이 2.292로 양(+의 값을 가지므로 상태벡터에 첨가한다.

그리고, 2 번째 행, 3번째 행은 새로운 원소가 첨가되었을 때 DIC가 음(-)의 값이므로 상태벡터에서 제외된다.

따라서 최종 상태벡터는 $\{X1_t, YT_t, X1_{t+1|t}, YT_{t+2|t}\}$ 이다.

Table 2. Canonical correlations analysis

State vector	Correlations	Infor.	Chisq	DF.
X1(T:T),YT(T:T),X1(T+1:T)	1.0000 1.0000 0.4291	8.891578	16.48455	4
X1(T:T),YT(T:T),X1(T+1:T),YT(T+1:T)	1.0000 1.0000 0.4325 0.2237	-1.73726	4.185703	3
X1(T:T),YT(T:T),X1(T+1:T),X1(T+2:T)	1.0000 1.0000 0.5029 0.1136	-4.92211	1.058408	3

선택된 상태벡터 공간을 이용하여 F, G의 최종 추정결과를 각 모수 추정치의 T-검정 결과 F(2,1), F(2,2), F(3,2)가 유의하지 못하므로 0으로 제한하였다. 다음은 상태벡터공간의 구성요소들의 설명한 것이다.

구성요소

$X1_t$, 현재 $X1$ 의 값 (월 평균 수온)

YT_t , 현재 YT 의 값 (월 평균 용존산 (DO))

$X1_{t+1|t}$, $X1$ 의 원점 t로부터 한 단계 앞의 예측치

상태벡터 공간들의 구성요소를 이용하여 우도함수가 최대가 되는 F, G를 구한 최종 상태공간 모형은 다음과 같다.

$$\begin{bmatrix} X1_t \\ YT_t \\ X1_{t+1|t} \end{bmatrix} = F \cdot \begin{bmatrix} X1_{t-1} \\ YT_{t-1} \\ X1_{t|t-1} \end{bmatrix} + G \cdot \begin{bmatrix} \epsilon_t^{X1} \\ \epsilon_t^{YT} \end{bmatrix} \quad (4.4)$$

F =

$$\begin{bmatrix} 0 & 0 & 1 \\ 0.480 & 0.210 & 0.000 \\ 0.011 & 0.000 & 0.271 \end{bmatrix}$$

G =

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -0.575 & 0.200 \end{bmatrix}$$

상태공간 모형의 식을 풀면 다음과 같다.

$$X_t = X_{t|t-1} + \epsilon_t^x$$

$$Y_t = 0.48 X_{t-1} + 0.21 Y_{t-1} + \epsilon_t^y$$

$$X_{t+1|t}$$

$$= 0.011X_{t-1} + 0.271X_{t|t-1} - 0.575 \epsilon_t^x + 0.2 \epsilon_t^y$$

$$X_{t|t-1} = 0.011X_{t-2} + 0.271X_{t-1|t-2}$$

$$- 0.575 \epsilon_{t-1}^y + 0.2 \epsilon_{t-1}^y$$

$$X_{t-1} = X_{t-1|t-2} + \epsilon_{t-1}^x \rightarrow X_{t-1|t-2} = X_{t-1} - \epsilon_{t-1}^x$$

$$X_t = 0.011X_{t-2} + 0.271X_{t-1|t-2}$$

$$- 0.575 \epsilon_{t-1}^x + 0.2 \epsilon_{t-1}^x$$

$$X_t = 0.011X_{t-2} + 0.271X_{t-1} - 0.271\varepsilon_{t-1}^x + 0.2\varepsilon_{t-1}^y \quad (4.5)$$

상태공간화한 식(4.5)을 벡터 ARMA(p,q)로 전환시키면 식(4.5)을 다음과 같이 나타낼 수 있다. 이 식은 VARMA(2,2)모형으로 추정된다.

$$X_t = 0.271X_{t-1} + 0.011X_{t-2} - 0.846\varepsilon_{t-1}^x + 0.2\varepsilon_{t-1}^y$$

$$Y_t = 0.48X_{t-1} + 0.21Y_{t-1} + \varepsilon_t^y$$

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} 0.271 & 0 \\ 0.48 & 0.21 \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} 0.011 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t^x \\ \varepsilon_t^y \end{bmatrix} + \begin{bmatrix} -0.846 & 0.2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{t-1}^x \\ \varepsilon_{t-1}^y \end{bmatrix} \quad (4.6)$$

설정된 모형이 적절한지에 대한 검증단계는 본 논문에서는 추정된 잔차를 분석하여 잔차가 백색잡음 과정을 갖는다면 적합한 모형이라고 할 수 있다.

최종 선택된 상태공간 모형 식(4.5)과 VARMA 모형 식(4.6)을 이용하여 모형에 의한 계산값과 실측값의 잔차에 대한 ACF와 PACF를 Fig. 4에 나타내었다. 그림에서와 같이 모든 시차에서 유의한

값이 없으므로 평균이 0이고 분산을 갖는 백색잡음 과정을 따른다고 할 수 있겠다. 따라서 상태공간모형은 장래의 월 용존산소예측에 사용될 수 있는 것으로 판정되었다. 설정된 모형이 적절한지에 대한 검증단계는 본 논문에서는 추정된 잔차를 분석하여 잔차가 백색잡음 과정을 갖는다면 적합한 모형이라고 할 수 있다.

3. 월 용존산소(DO) 검증 및 수질예측

달천지점의 월 용존산소(DO) 예측을 위하여 분석된 승법ARMA 모형과 상태공간 모형으로부터 월 용존산소(DO)를 예측하였다.

단변량 시계열 모형인 승법ARIMA(2,1,1)(1,1,1)₁₂ 모형에 의한 예측은 식(4.1)을 사용하여 예측할 수 있다. 그리고 다변량 시계열 모형인 상태공간 모형은 식(4.6)으로 예측할 수 있다.

Fig. 5.는 단변량 시계열 모형인 승법 ARIMA (2,1,1)(1,1,1)₁₂모형에 의한 계산값과 실측값을 비교하여 재현성을 보여주는 것으로 비교적 잘 일치하는 것으로 나타났다.

다변량시계열 모형인 상태공간모형에 의한 계산값과 실측값을 Fig. 6.에 나타냈다.

그림에서와 같이 상태공간 모형에 의한 계산값과 실측값들이 재현성이 매우 우수한 것을 나타냈다.

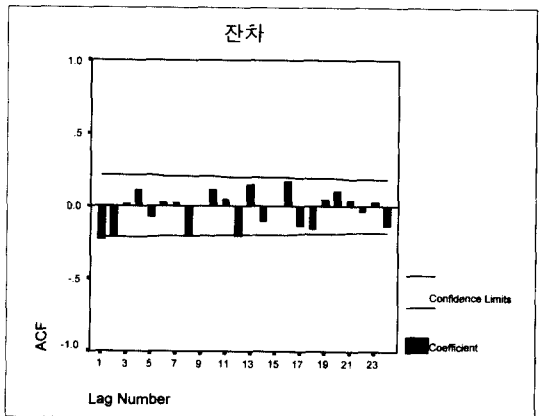
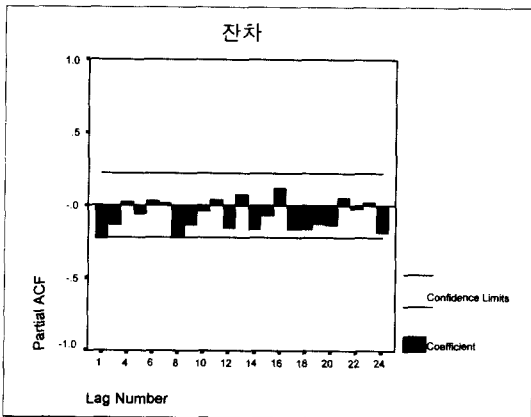


Fig. 4. ACF and PACF of residual estimated by State-space model(DO)

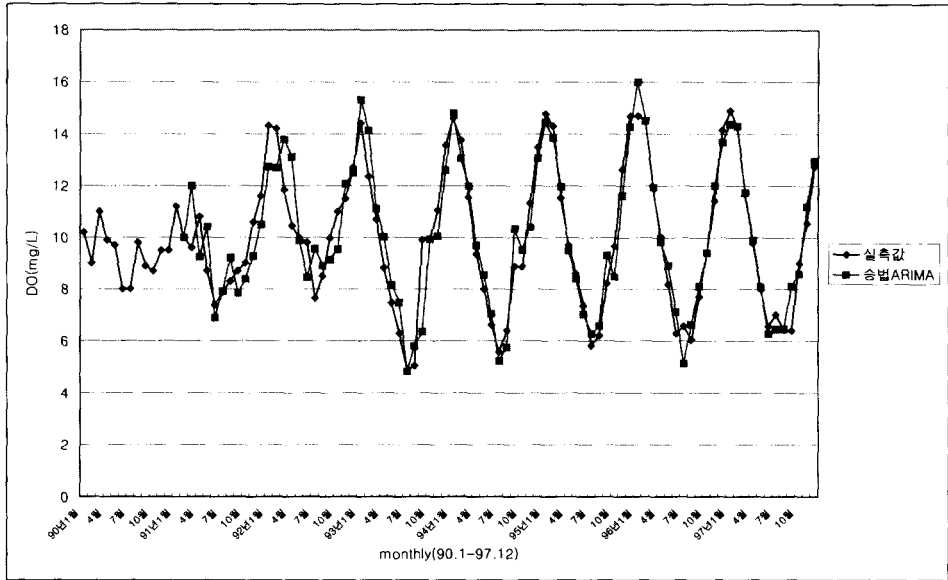


Fig. 5. Comparison between observed and estimated values by multiplicative ARIMA(2,1,1)(1,1,1)₁₂ model in Dalchun station

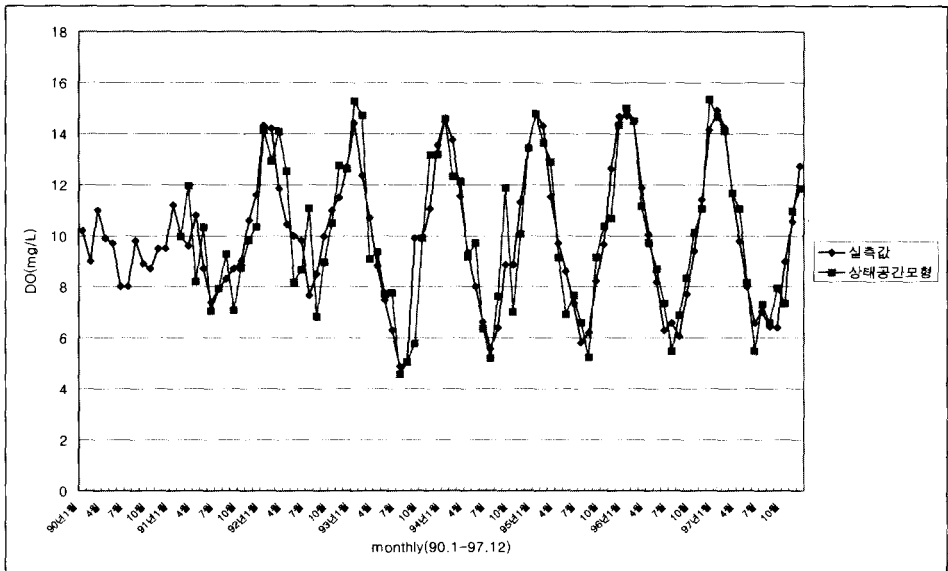


Fig. 6. Comparison between observed and estimated values by State space model in Dalchun station

달천지점의 월 평균 용존산소 예측을 위하여 분석된 승법 $ARIMA(2,1,1)(1,1,1)_{12}$ 모형과 상태공간모형을 사용하여 1998년 1월부터 1999년 10월까지 예측하여 예측값과 실측값을 비교 검증하였다

Table 3. Characteristics of statistics on water quality by each model

	Observed	SARIMA	State-space model
Average	8.70	7.94	9.29
Standard deviation	2.75	3.17	3.01
Skewness	0.60	0.51	0.62
Correlation	-	0.93	0.97

월 용존산소 예측모형인 단변량 시계열 모형 승법 ARIMA(2,1,1)(1,1,1)₁₂모형에 의한 예측값과 다변량 시계열 모형인 상태공간 모형의 예측값과 실측값을 Fig. 7.에 나타내었다. 그리고 통계특성은 Table 3.에 나타내었다.

실측값과 승법 ARIMA(2,1,1)(1,1,1)₁₂모형, 상태공간 모형에 의한 예측값의 평균은 각각 8.70, 7.94, 9.29mg/L 로 나타났고, 표준편차는 각각 2.75, 3.17, 3.01로 나타났다.

실측값에 대한 승법 ARIMA(2,1,1)(1,1,1)₁₂모형, 상태공간모형의 예측값의 상관계수는 각각 0.93과 0.97로 나타났다.

따라서 승법 ARIMA(2,1,1)(1,1,1)₁₂모형 보다 상태공간 모형의 예측력이 약간 좋은 것으로 나타났다. 따라서 이들 모형을 사용하여 달천 지점의 월 용존산소 장래 예측에 사용될 수 있는 것으로 나타났다.

상태공간 모형에 의한 장래의 DO농도의 예측값은 1999년 10월부터 2002년 12월까지 월 농도를 예측하여 Fig. 8.에 나타내었다.

그림에서와 같이 월 평균 용존산소의 예측결과 일정한 계절 주기를 따르고 이월효과가 큰 특성을 나타내고있으며 이는 수온에 대한 영향이 큰 것으로 사료된다. 그리고 연도가 증가할수록 DO값이 낮아지는 것으로 예측되었다.

V. 결 론

본 연구는 달천강 유역의 하류인 충주시 상수원 취수지점에서 월 수온자료와 월 용존산소(DO) 자료를 사용하여 승법 ARIMA 모형과 상태공간 모형을 적용하였다. 월 용존산소(DO)의 장·단기 예측이 가능하도록 하여 다음과 같은 결과를 얻을 수 있었다.

- (1) 월 용존산소(DO)와 월 수온 자료는 이월효과

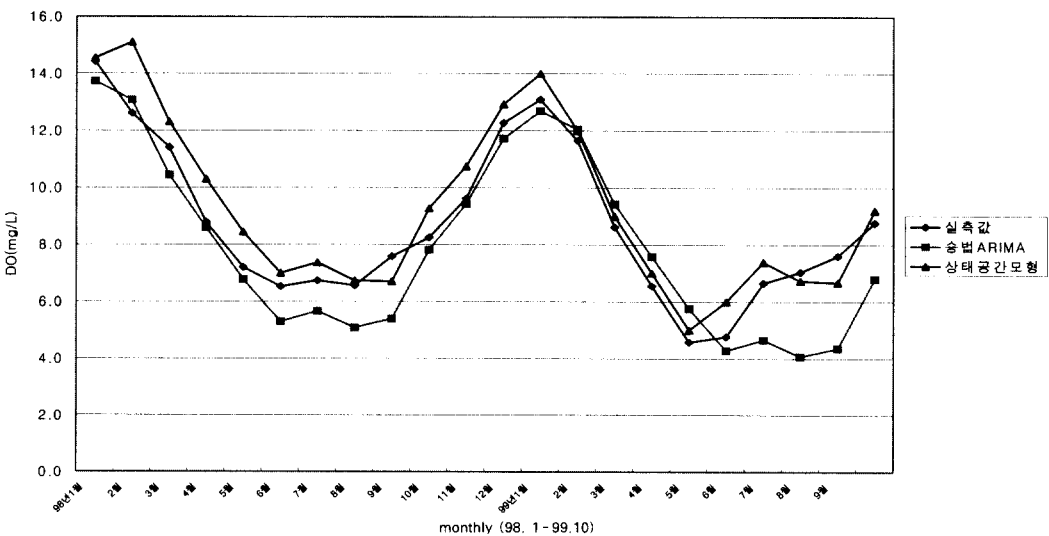


Fig. 7. Comparison of observed and forecasting values by state space model multiplicative ARIMA(2,1,1)(1,1,1)₁₂ (DO)

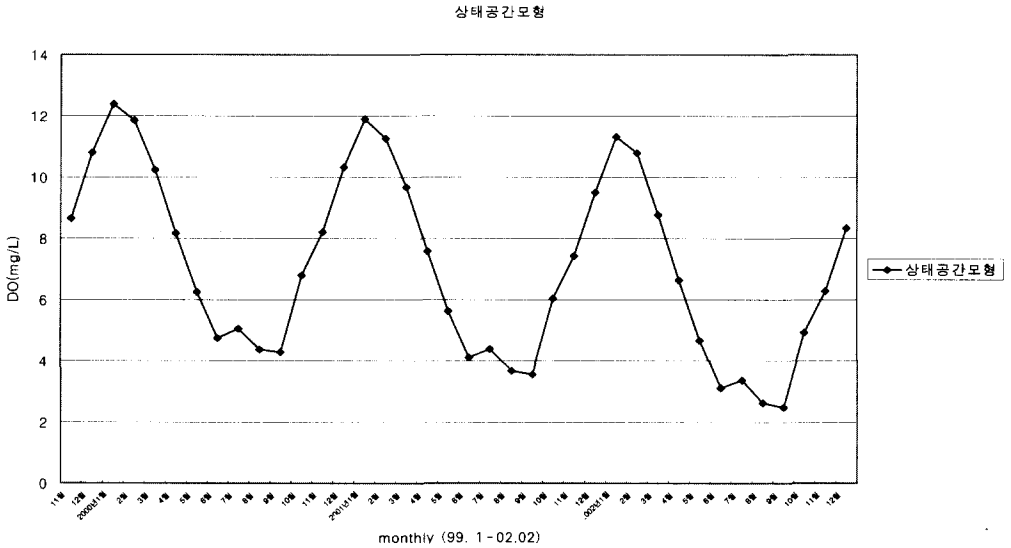


Fig. 8. Forecasting concentration of DO using a state space model

가 강한 지속성과 12월의 주기성을 갖고 있으며, 이를 분석한 결과 월 용존산소(DO)와 월 수온은 동일하게 승법 ARIMA(2,1,1)(1,1,1)₁₂ 모형으로 선택할 수 있었다.

- (2) 유역의 시스템에서 월 수온과 월 용존산소(DO)를 입력치로 하여 월 용존산소(DO)를 출력치로 하는 상태공간모형을 설정할 수 있었다.
- (3) 승법 ARIMA모형의 예측값과 실측값 그리고 상태공간 모형의 예측값과 실측값에 대한 각각의 상관계수 R은 0.93, 0.97로 나타났으며 실측값에 대한 표준편차는 각각 3.17, 3.01로 나타났다.
- (4) 월 용존산소(DO) 예측을 위한 승법 ARIMA 모형과 상태공간 모형을 비교한 결과 다변량 시계열 모형인 상태공간 모형이 예측의 신빙성이 우수하였다.

위 결과에 의하여 상태공간 모형을 사용하여 달린 취수지점에서 장기 월 용존산소 예측에 유용할 것으로 판단된다.

참고문헌

1. 김수원, “도시용수 계획을 위한 물 수요예측에 관한 연구”, 박사학위 논문, 영남대학교, 1985.
2. 이경훈, 박성천, 문병석, “상수도의 총수량 및 저수지 용량변화에 관한 연구”, 대한상하수도학회, 제9권, 제2호, pp. 118~126, 1994a.
3. 이경훈, 이삼노, 문병석, “상수도1일 총수량 변화와 확률분산 모형에 관한 연구”, 제36회 수공학 연구발표회 논문집, 한국수문학회, pp. 405~410, 1994b.
4. 이준호, “ARIMA 모형을 이용한 한강의 DO 변화분석”, 석사학위논문, 서울대학교, 1991.
5. 최병선, “SAS를 이용한 시계열분석시리즈4, 다변량 시계열분석”, 세경사, 1995.
6. Box, G. E. P. and Jenkins., “G. M. Time series analysis-forecasting and control”, revised edition, Holden Day, 1976.
7. “SAS/ETS user’s guide”, SAS Institute Inc., USA, 1993.