



Development of salinity simulation using a hierarchical bayesian ARX model

Kim, Hojun^a · Shin, Choong Hun^b · Kim, Tae-Woong^c · Kwon, Hyun-Han^{d*}

^aMaster Course, Department of Civil and Environmental Engineering, Sejong University, Seoul, Korea

^bAssociate Researcher, Rural Research Institute, Korea Rural Community Corporation, Ansan, Korea

^cProfessor, Department of Civil and Environmental Engineering, Hanyang University, Ansan, Korea

^dProfessor, Department of Civil and Environmental Engineering, Sejong University, Seoul, Korea

Paper number: 20-021

Received: 11 March 2020; Revised: 26 May 2020; Accepted: 26 May 2020

Abstract

The development of agricultural land at Saemangeum has required a significant increase in agricultural water use. It has been well acknowledged that salinity plays a critical role in the farming system. Therefore, a systematic study in salinity is necessary to better manage agricultural water. This study aims to develop a stochastic salinity simulation model that simultaneously simulates salinities obtained from different layers. More specifically, this study proposed a two-stage Autoregressive Exogeneous (ARX) model within a hierarchical Bayesian modeling framework. We derived posterior distributions of model parameters and further used them to obtain the predictive posterior distribution for salinities at three different layers. Here, the BIC values are used and compared to determine the optimal model from a set of candidate models. A detailed discussion of the model is provided.

Keywords: Hierarchical bayesian, Autoregressive model, ARX model, Salinity, Saemangeum

계층적 베이시안 ARX 모형을 활용한 염분모의기법 개발

김호준^a · 신충훈^b · 김태웅^c · 권현한^{d*}

^a세종대학교 건설환경공학과 석사과정, ^b한국농어촌공사 농어촌연구원 주임연구원,

^c한양대학교(ERICA) 건설환경공학과 교수, ^d세종대학교 건설환경공학과 교수

요 지

새만금 농업단지가 조성됨에 따라 농업용수 공급이 요구되며, 농업적 측면에서 염분은 작물 재배시 생육에 영향을 미치는 항목으로 농업용수 공급 시 철저한 관리가 요구된다. 따라서 농작물에 영향을 미치지 않는 농업용수 공급을 위해 염분계측을 통한 체계적인 농업용수 관리가 필요하다. 본 연구에서는 새만금호내에 관측되는 염분 시계열 자료를 모의하기 위해서 자기회귀모형을 기반으로 한 Two-Stage ARX 모형을 개발하였다. 층별로 나뉜 염분자료를 계층적 Bayesian 기법을 활용하여 매개변수를 확률분포형으로 추정하였으며 염분모의의 불확실성을 제시하였다. 최적 모형을 선정하기 위해서 통계적 지표인 BIC값을 이용하였으며, 최종적으로 선정된 모형을 통해 양수장 인근 수역의 염분 모의 결과를 제시하였다.

핵심용어: 계층적 베이시안, 자기회귀모형, ARX 모델, 염분, 새만금유역

1. 서 론

새만금 간척사업의 진행으로 대규모 농업단지가 건설될 예정이며 그에 따른 대량의 농업용수 공급이 요구된다. 환경

부에서는 농업용수의 수질환경기준을 제시하고 있으며, 공급용수는 농업용수 수질기준을 만족해야 한다. 새만금 호의 수질에 대한 연구가 진행되고 있으며, 수질 및 염분에 대한 모니터링, 모의 및 예측 등을 바탕으로 이루어지고 있다. Suh and Lee (2008)은 방조제 완공 후 변화된 수리특성의 영향에 따른 새만금 호내의 수질 변화 모의를 제시하였다. 모의 결과 새만금 호내 저층에서 DO농도가 급격하게 감소하는 빈산소 현상

*Corresponding Author. Tel: +82-2-3408-3726
E-mail: hkwon@sejong.ac.kr (H.-H. Kwon)

이 발생하였으며, 그 원인으로는 배수갑문만을 통한 해수의 교환율 감소에 따라 수온 및 염도성층에 의해 표층과 저층 간의 물질교환 작용이 감소와 퇴적물 내 유기물들의 분해로 인하여 산소요구량(SOD; Sediment Oxygen Demand) 증가를 요인으로 판단하였다. Jeong and Yang (2015)은 새만금호의 장기적인 수질변화를 위하여 모니터링을 실시하였으며, 방조제 끝물막이 직후의 수질인자의 변화 결과를 제시하였다. 수질 모니터링 결과 하계에 강 하류부에서 강한 염분 성층과 표층에서의 저산소 상태를 관측하였으며, 이는 퇴적물로부터 용출을 유발하여 수질악화 가속화의 발생 가능성을 제시하였다. 방조제 끝물막이 이후 해수 교환율 감소와 만경·동진강의 담수 유입의 영향과 함께 염분농도가 점차적으로 감소하였고, 이후 장기간 염분은 감소된 상태를 유지하였다. 그중 염분은 농업적 측면에서 작물의 성장을 방해하여 생산성을 저하시키는 결과를 초래하기에 농업용수 취수에 있어서 중요하게 고려해야 할 요소이다(Bauder *et al.*, 2011). 만경·동진수역에 양수장을 운영하여 농업용수를 공급할 예정이지만, 기간에 따른 염분농도의 불규칙한 변화로 인하여 취수가 어려운 실정이다(ME, 2011; Jeong and Yang, 2015; MAFRA, 2017). 수질기준에 적합한 농업용수 공급을 위하여 양수장 인근에서의 염분계측자료 및 영향인자를 고려한 염분 모의가 필요하며, 이를 활용하여 양수장 운영방안을 제시할 수 있다.

현재 염분의 영향으로 인한 피해를 방지하기 위해 염분에 대한 연구가 수행되고 있지만, 국내 연구에서는 염분에 대한 집중적인 연구보다는 수질인자들을 종합적으로 다루며 결정론적 모형과 추계학적 모형을 활용한 연구가 다수 진행된 바 있다. 결정론적 모형은 정량적인 입력에 따른 결과를 표출해주는 과정을 물리적인 식을 통해 표현되는 모형을 의미하며, 주로 3차원 수리모형인 Environmental Fluid Dynamics Code (EFDC) 등이 활용되고 있다. Cho *et al.* (2018)은 낙동강 하구 농업용수 취수원 방어를 위해 EFDC를 활용한 염수침입거리 산정을 통해 최적 수문운영을 제시하였고, Suh and Lee (2008)는 새만금 방조제 공사 이후 배수갑문을 통한 해수유통이 수질에 미치는 영향을 분석하기 위해 EFDC 모형을 활용한 수치모의를 실시하는 등 결정론적 모형을 활용한 연구를 수행하였다. 추계학적 모형은 집단의 특성을 통계적으로 표현하여 모수를 추정하는 방법을 뜻하며, 장기간의 관측자료로 구축된 모형의 모의 발생방법을 사용하여 복잡한 수자원 시스템의 계획, 설계 등의 입력자료로 제시되고 있다. 대표적인 추계학적 모형으로는 Monte-Carlo 모형, spectral 해석모형, Box-Jenkins의 시계열 모형, Kalman filter 모형 등이 있다. Cho *et al.* (2008)은 아산만 연안의 담수 영향범위 추정을 위하여 계절별 염분

과 담수유출량과의 관계를 파악하는 선형회귀식과 상관관계를 제시하였으며, 염분농도를 기준으로 담수 확산 범위를 분석하였다. Park (2002)는 수질예측모형을 위해 시계열 분석인 Autoregressive Integrated Moving Average (ARIMA) 기법을 활용하였으며, 한강수계를 대상으로 수질데이터를 분석하여 계절적 변화 양상과 모형 적용성을 평가하였다. 국외에서는 추계학적 모형을 활용한 연구가 진행되었다. Marshall III *et al.*, (2004)은 풍속과 조위를 고려한 Multivariate Linear Regression (MLR) 기반의 염분모의모형을 제시하였고, Qiu and Wan (2013)은 하천 유입량, 강우량, 조위 등을 독립변수로 하는 자기상관회귀모형을 활용하여 염분모의를 수행한 결과를 3D 수치해석 모형과 비교하였다. 그 결과, 두 모형은 유사한 결과를 보였지만, 3D 수치해석은 통계 모형에 비해 비용과 시간이 크게 요구되는 등 효율성 측면에서 불리한 측면이 있다. 이처럼 효율적 측면에서 장점이 있는 통계적 모형을 활용하면 취수에 있어서 단기적으로 모의 및 예측이 가능하다. 국외에서는 다양한 통계적 모형의 적용을 통해 염분 모의 및 예측에 대한 연구가 진행되고 있지만, 국내에서는 염분보다는 수질인자에 초점을 두어 연구되고 있어 추계학적 모형의 개발 및 적용이 제한적으로 이루어지고 있다. 본 연구에서는 새만금호내의 염분 관측자료를 이용한 추계학적 기반 염분모의 모형을 구축하여 계절적으로 나타나는 염분 패턴을 분석하였다.

본 연구 수행과정은 다음과 같이 진행되었다. 새만금호내 월 염분자료와 강수량 자료를 시계열 형태로 구축한 후 자료의 전처리를 위해 Log-scale 변환 및 정규화를 수행하였다. 자기회귀모형인 Autoregressive (AR) 모형에 독립변수가 추가된 Autoregressive Exogenous (ARX) 모형을 기반으로 영향인자(독립변수, X)와 취수장 근처의 염분(종속변수, Y)과의 관계식을 나타내는 염분모의모형을 구축하였다. 수심에 따라 나뉜 염분 시계열자료들이 통계적으로 유의한 특성을 가짐으로 계층적 Bayesian 기법을 통해 매개변수의 사후분포를 추정하였다. 통계적으로 모형을 판단하는 지표인 Bayesian Information Criterion (BIC)을 통해 기상인자 포함 유무에 따라 나눈 모형을 비교하여 최적 모형을 선정하여 제시하였다. 본 논문의 구성은 다음과 같다. 2장에서는 새만금유역의 염분자료를 대상으로 본 연구에서 제안하는 Two-Stage ARX 모형의 개념 및 해석 절차를 서술하였고 3장에서는 최적모형 선정 및 모형 검증을 수행하였다. 4장에서는 해석 결과에 대한 고찰 및 향후 연구방향에 대하여 서술하였다.

2. 연구방법

본 연구에서는 새만금호내에 위치한 2011-2017년도 염분 모니터링 자료와 부안 기상관측소의 강수량 자료를 활용하였으며, 계절에 따라 변화가 큰 염분을 재현하기 위하여 임계점을 기준으로 두 가지로 분류한 ARX 모형에 대한 서술과 통계적으로 유사한 특성을 지닌 층별 염분자료에 적용한 계층적 Bayesian 기법을 기술하였다.

2.1 대상자료

새만금의 유역은 크게 만경강과 동진강유역으로 나뉘며 만경강과 동진강은 새만금호내로 유입되어 신시·가력배수갑문을 통해 서해로 유출된다. 만경강은 유역면적 1,527.1 km², 유로연장 77.4km의 국가하천으로 전라북도 완주군 동상면 원등산(EL. 713 m)에서 발원하여 지류와 합류한 뒤 새만금호로 유입되고, 동진강은 유역면적 1,136.2 km², 유로연장 51.0 km의 국가하천으로 전라북도 정읍시 산외면 상두리 국사봉(EL.543 m)에서 발원하여 새만금호로 유입된다. 새만금유역의 염분은 2010년 말 이후 관리수위 저하(-1.6 m)로 인한 해수 유통량 감소와 하천수 유입으로 염분농도가 점차 감소하였으며, 관리수위 조정 이후 주기적으로 새만금 내부개발 사업을 위하여 한국농어촌공사 새만금사업단에서 수질 및 염분 경향을 파악하고 있다. 향후 농업용수 공급 및 수질관리 대책 수립을 위한 기초자료로써 활용하기 위해 새만금호내 염분 및 상시수질 모니터링을 수행하고 있다.

본 연구에서는 한국농어촌공사 새만금사업단에서 2011-2017년에 측정된 자료를 최대수심에 따라 분류된 층별(표층, 중층, 저층) 월평균 염분 시계열자료를 활용하였다. 분류된 층은 최대수심을 기준으로 표면에서의 1/3 지점까지 표층, 1/3 지점에서 2/3 지점까지 중층, 그 아래를 저층으로 구분하였으며, 결측치가 존재하여 기존 자료의 월평균 염분을 산정하여 그 값으로 대체하였다. 동진강 하류부에 위치한 농업용수 취

수 예정지인 동진양수장(동진지점)과 배수갑문 근처에 위치하여 해수의 영향을 직접적으로 받는 상시모니터링 지점(D4)을 분석 지점으로 선정하였다. 기상인자는 새만금유역에 가깝게 위치한 기상청산하 부안 종관기상관측소(ASOS)의 월 강수량자료를 사용하였다. Fig. 1은 본 연구에서 활용한 새만금유역내 동진지점, 상시모니터링(D4), 부안 기상관측소를 도시한 결과를 나타내었다.

염분은 유량의 영향을 받으며, 특히 하천 유량의 급격한 변화는 염분에 영향을 크게 미친다(Wang *et al.*, 2011). 우리나라의 계절 특성상 여름철에 집중적으로 강우가 발생하며 따라서 염분은 계절적인 주기를 가진다. Fig. 2는 동진지점과 상시모니터링 지점(D4)에서 관측된 층별(표층, 중층, 저층) 월평균 염분자료를 나타내었다. Fig. 2에서 확인할 수 있듯이, 담수 유입량이 큰 여름철에는 염분이 작게 관측되며 유입량이 적은 겨울철에는 염분이 크게 나타난다. 또한 배수갑문을 통해 유입된 해수로 인한 동진강 하류부까지 염분전이가 지체됨을 확인할 수 있다. 본 연구에서는 새만금호내에서 동진강 상류측으로 이어지는 염분의 전이정도를 추계학적 모형을 통해 모의하는 데 목적이 있으며, 자료가 가지는 비정상성을 고려

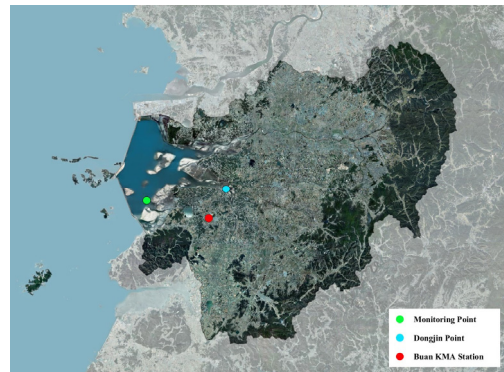


Fig. 1. The map showing Mangyeong-river and Dongjin-river watershed along with salinity monitoring stations and weather station

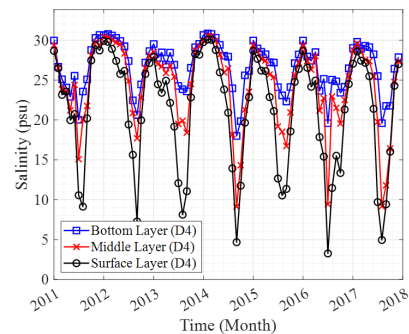
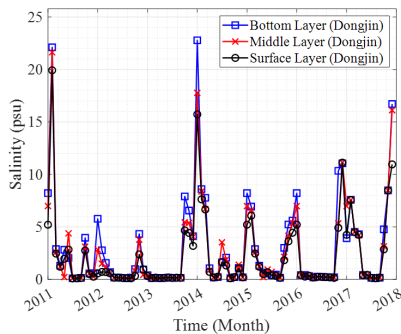


Fig. 2. A graphical representation of average monthly salinity for each layer at Dongjin and Monitoring stations (D4)

하기 위하여 외부인자를 고려하였다. 이를 위하여 배수갑문 주변의 비정상성(감소 경향성)을 가지는 염분농도를 독립변수로 활용하여 간접적으로 동진강 상류의 염분변화 모의가 가능토록 하였다.

2.2 Two-Stage ARX 모형

자기회귀 모형은 생물학, 기상학, 경제학 등 많은 분야에서 활용되고 있으며, 특히 수문학 분야에서 시계열 모의 및 예측을 위해 활용되고 있다(Lettenmaier and Wood, 1993; Salas, 1980). 유량과 유사하게 염분 시계열자료는 자기상관성이 존재하며, 시계열 자료 분석에 있어 중요한 역할을 하므로 본 연구에서는 이를 활용한 자기회귀모형을 제안하고자 한다. 자기회귀는 하나의 시계열자료를 시간 차이를 두어 종속변수와 독립변수로 표현되는 관계식을 의미한다. 일반적으로 연속적인 일련의 시계열자료들이 상관되어 있을 때 상관계수를 산정하여 선형적으로 상관관계를 나타낸다. 마찬가지로 하나의 시계열자료에 대해서 p시점 이전의 값과 현재의 값과의 상관계수를 산정하여 자기상관관계를 나타낼 수 있으며, 통계적으로 유의하다면 현재의 값은 과거에 해당한 값의 영향을 받는 것을 의미한다. 시계열자료 y_t 가 존재할 때, 자기상관계수는 Eq. (1)과 같이 정의된다.

$$\hat{\rho}_p = \frac{\sum_{t=1}^T (y_t - \bar{y})(y_{t-p} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2} \quad (1)$$

p계 자기회귀모형(AR(p))은 p시점에 해당하는 Y값이 과거 기간((t-1), (t-2), ...)에 해당하는 Y값에 영향을 받는 모형을 의미하며, 종속변수(Y_t)와 독립변수(Y_{t-1}, \dots, Y_{t-p})로 표현되는 선형회귀모형이며, Eq. (2)와 같이 표현된다.

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \epsilon_t \quad (2)$$

p가 1인 모형을 AR(1)모형이라고 하며 현재값과 이전 기간 (t-1)의 자기 자신의 값만으로 이루어진 모형이다. ARX 모형은 기존의 자기회귀모형인 AR모형에 외부요인인 독립변수 X를 추가한 모형으로 종속변수와와의 관계를 외부정보가 추가된 관계식을 통해 보여준다. Eq. (3)은 현재 $y_{(t-1)}$ 부터 p_y 단계 이전의 $y_{(t-p_y)}$ 으로 구성된 선형조합과 현재 $x_{(t-1)}$ 에서 p_x 단계 이전의 $x_{(t-p_x)}$ 까지의 선형조합과 오차 ϵ_t 로 이루어져 있다.

$$y_t = \beta_1 x_{(t-1)} + \beta_2 x_{(t-2)} + \dots + \beta_{n_x} x_{(t-p_x)} + \alpha_1 y_{(t-1)} + \alpha_2 y_{(t-2)} + \dots + \alpha_{n_y} y_{(t-p_y)} + \epsilon_t \quad (3)$$

본 연구에서는 동진지점 염분자료의 자기상관성을 확인하기 위해 Fig. 3에 자기상관도를 제시하였으며, 이를 통해 자기상관성을 확인하였으며 이전 기간(t-1) 염분의 자기상관계수가 $\pm \frac{1.96}{\sqrt{T}}$ 보다 크기에 통계적으로 유의한 것으로 판단하였다. 이는 이전 기간(t-1) 염분이 현재 염분에 영향을 미치는 것을 의미하며 서로 간의 상관관계가 존재한다는 것을 나타낸다. Fig. 3에서 이전 기간(t-1)을 제외한 염분은 5% 유의수준에서 자기상관이 존재하지 않는다고 판단된다. 동진 지점의 이전 기간의 염분 ($Y_{(t-1)}$)이 동진 지점 염분(Y_t)에 대한 표현이 가능하지만, 염분모의를 수행하기엔 예측성 및 시간적인 한계가 발생한다. 따라서 하류부에서 비교적 멀리 위치한 D4 지점의 염분을 독립변수(X_t)로 선정하였다. 추가로 기상인자의 영향을 확인하고자 수문기상학적 자료인 월 강수량을 선정하였다.

강우-유출 관계에서 강우와 유출간의 시간차를 고려하기 위하여 본 연구에서는 1개월 지체시간을 갖는 월강우량 ($R_{(t-1)}$)을 독립변수로 고려하였다. Table 1은 동진지점 염분과 독립변수간의 층별 상관계수를 나타낸다. ARX 모형을 구성하는 데 있어서 지체시간 결정은 자기상관도(autocorrelation) 및 편자기상관성(partial autocorrelation)를 통해 결정하였다.

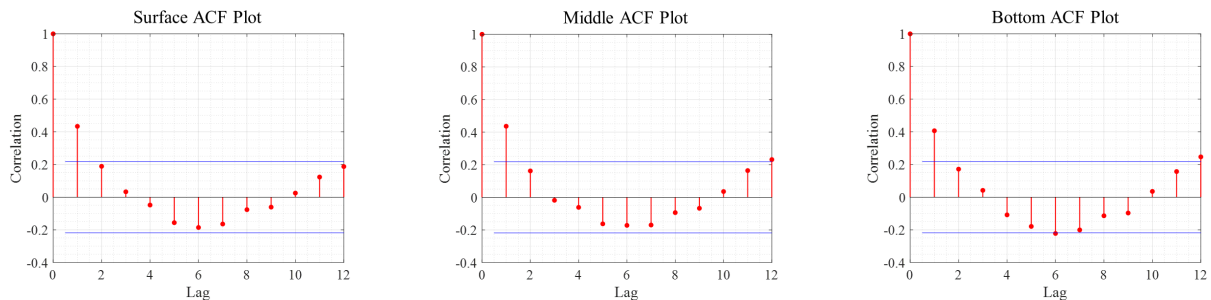


Fig. 3. Autocorrelation function of average monthly salinity for each layer at Dongjin station

Table 1. A comparison of coefficient between salinity and influencing factors

Layer	1 month-lag salinity ($Y_{(t-1)}$)	D4 Station salinity (X_t)	1 month-lag rainfall ($R_{(t-1)}$)
Surface	0.4523	0.4054	-0.3390
Middle	0.4714	0.3383	-0.3485
Bottom	0.4346	0.2929	-0.3316

염분농도는 계절적으로 편차가 크며, 이를 하나의 모형으로 설명하는 데 한계가 있다. 이러한 점에서 본 연구에서는 상대적으로 큰 염분상태를 고려한 모형을 구축하기 위하여 임계값(threshold) θ 을 기준으로 나누는 후 이를 종속변수와 비교할 수 있는 모형을 제안하고자 한다. 염분농도에 따라 자동으로 구간을 분리할 수 있도록 지시함수(indicator) I 를 고려하여 모형을 구축하였다. 즉, ARX 모형의 종속변수(Y_t)를 기준이 되는 임계점 θ 에 따라 두 가지 경우로 나누었기에 Two-Stage ARX 모형으로 명명하였다. 기상인자 포함여부에 따라 모형을 두 가지로 구분하였으며 Eq. (4a)은 기상인자를 포함하지 않고 Y_{t-1} 과 X_t 로 구성된 모형이고 Eq. (4b)은 Eq. (4a)에 추가적인 독립변수로 월강우량을 포함한 모형을 나타낸다.

$$Y_t = \begin{cases} (\alpha_1 Y_{t-1} + \beta_1 X_t + \gamma_1) I(\theta < Y_t) \\ (\alpha_2 Y_{t-1} + \beta_2 X_t + \gamma_2) I(\theta \geq Y_t) \end{cases} \quad (4a)$$

$$Y_t = \begin{cases} (\alpha_1 Y_{t-1} + \beta_1 X_t + \gamma_1 R_{t-1} + \delta_1) I(\theta < Y_t) \\ (\alpha_2 Y_{t-1} + \beta_2 X_t + \gamma_2 R_{t-1} + \delta_2) I(\theta \geq Y_t) \end{cases} \quad (4b)$$

염분 시계열 자료 특성상 계절적인 주기와 복합적인 요인으로 변동하는 경우와 평년값을 벗어나는 극치사상이 존재하기에 염분 시계열 자료를 하나의 관계식으로 표현하는 것이 어려워진다. 서로 다른 양상을 보이는 시계열 자료를 분리하기 위해서 임계점을 설정하였으며, 이렇게 분리된 모형을 통해 자료의 변동점 및 극치값을 재현할 수 있다.

2.3 계층적 Bayesian 모형

본 연구에서 사용된 층별로 관측된 염분자료는 통계적으로 유사하기에 계층적 Bayesian 기법을 활용하여 매개변수 추정을 하고자 한다. 계층적 Bayesian 기법은 유사한 통계적 특성을 가지는 자료집단들이 정보를 서로 교환함으로써 추정되는 매개변수의 불확실성을 저감하도록 고안된 방법이다 (Kwon et al., 2013). 본 연구에서는 동질성이 확보된 염분자료를 부분적 통합(partial pooling)하여 모집단 분포로부터 추출된 매개변수가 층별로 구분된 염분자료간의 의존성을 고려

한 매개변수를 산정할 수 있다. 추가적으로 매개변수의 불확실성을 정량화할 수 있도록 모형을 개발함으로써 추정되는 염분자료의 신뢰성을 개선할 수 있다.

Bayesian 추론기법은 매개변수 최적화 및 불확실성 정량화를 목적으로 두며 매개변수를 단일 값이 아닌 확률분포 형태로 부여되고 결합확률분포(joint probability distribution)는 사전분포(prior distribution) $p(\theta)$ 와 우도(likelihood) $p(y|\theta)$ 의 곱으로 표현된다는 베이즈 정리(Eq. (5))를 기반으로 사후분포(posterior distribution) $p(\theta|y)$ 를 추론한다.

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} = \frac{p(\theta)p(y|\theta)}{p(y)} \propto p(y|\theta)p(\theta) \quad (5)$$

본 연구에서는 종속변수인 염분자료의 확률분포형을 대수 정규분포로 가정하였으며 계층적 Bayesian 모형은 Eq. (6)과 같이 표현된다.

$$Y_{t,n} \sim N(\mu_{t,s} | ((\alpha_{1,s} Y_{t-1} + \beta_{1,s} X_t + \gamma_{1,s}) I(\theta_s < Y_t) + (\alpha_{2,s} Y_{t-1} + \beta_{2,s} X_t + \gamma_{2,s}) I(\theta_s \geq Y_t)), \sigma_{t,s}^2) \quad s = 1, 2, 3 \quad (6a)$$

$$\alpha_{1,s} \sim N(\mu_{\alpha_1}, \sigma_{\alpha_1}^2), \beta_{1,s} \sim N(\mu_{\beta_1}, \sigma_{\beta_1}^2), \gamma_{1,s} \sim N(\mu_{\gamma_1}, \sigma_{\gamma_1}^2) \quad s = 1, 2, 3 \quad (6b)$$

$$\alpha_{2,s} \sim N(\mu_{\alpha_2}, \sigma_{\alpha_2}^2), \beta_{2,s} \sim N(\mu_{\beta_2}, \sigma_{\beta_2}^2), \gamma_{2,s} \sim N(\mu_{\gamma_2}, \sigma_{\gamma_2}^2) \quad s = 1, 2, 3 \quad (6c)$$

$$\theta_s \sim G(\theta_k, \theta_h) \quad s = 1, 2, 3 \quad (6d)$$

$$\sigma_{t,s} \sim G(\sigma_k, \sigma_\theta) \quad s = 1, 2, 3 \quad (6e)$$

여기서, $Y_{t,n}$ 는 벡터자료로서 세 개의 층으로 나누어진 동진 지점의 월 염분자료를 의미하며 $\alpha_{1,s}, \beta_{1,s}, \gamma_{1,s}, \alpha_{2,s}, \beta_{2,s}, \gamma_{2,s}, \theta_s, \sigma_{t,s}$ 는 정규분포를 사전분포를 갖는 회귀계수를 나타낸다. Eq. (6a)의 모든 매개변수는 확률분포 형태로 표현되며 $\alpha_{1,s}, \beta_{1,s}, \gamma_{1,s}, \alpha_{2,s}, \beta_{2,s}, \gamma_{2,s}, \theta_s, \sigma_{t,s}$ 은 사전분포를 가지며 정규분포와 감마분포로 추정되는 것으로 가정하였다. 8개의 매개변수 추정에 있어서 충분한 자료들이 존재하기에 Non-Informative 사전분포를 통해 수행하였다(Gelman et al., 2004; Gelman, 2006). 사후분포를 추정하기 위해서는 Eq. (5)의 사전분포와 우도함수를 통해 표현되며 계산된다. Eq. (5)에서 $\theta = [\alpha_{1,s}, \beta_{1,s}, \gamma_{1,s}, \alpha_{2,s}, \beta_{2,s}, \gamma_{2,s}, \theta_s, \sigma_{t,s}, \mu_{\alpha_1}, \sigma_{\alpha_1}, \mu_{\beta_1}, \sigma_{\beta_1}, \mu_{\gamma_1}, \sigma_{\gamma_1}, \mu_{\alpha_2}, \sigma_{\alpha_2},$

$\mu_{\beta_2}, \sigma_{\beta_2}, \mu_{\gamma_2}, \sigma_{\gamma_2}, \theta_k, \theta_h, \sigma_k, \sigma_\theta]$, $s = 1, 2, 3$ 으로서 s 는 세 개로 구분된 층을 나타낸다. 우도함수인 $p(y|\theta)$ 는 Eq. (7)과 같이 표현된다.

$$p(y|\theta) = \prod_{s=1}^3 \prod_{t=1}^n N(\mu_{t,s} | ((\alpha_{1,s} Y_{t-1} + \beta_{1,s} X_t + \gamma_{1,s}) I(\theta_s < Y_t) + (\alpha_{2,s} Y_{t-1} + \beta_{2,s} X_t + \gamma_{2,s}) I(\theta_s \geq Y_t)), \sigma_{t,s}^2) \quad (7)$$

Eqs. (6) and (7)에서 정의되는 값을 Eq. (5)에 대입시킴으로써 매개변수들의 사후분포를 Eq. (8)와 같이 추정할 수 있다.

$$p(\theta|y) = \prod_{s=1}^3 \prod_{t=1}^n N(\mu_{t,s} | ((\alpha_{1,s} Y_{t-1} + \beta_{1,s} X_t + \gamma_{1,s}) I(\theta_s < Y_t) + (\alpha_{2,s} Y_{t-1} + \beta_{2,s} X_t + \gamma_{2,s}) I(\theta_s \geq Y_t)), \sigma_{t,s}^2) \cdot N(\alpha_{1,s} | \mu_{\alpha_1}, \sigma_{\alpha_1}^2) \cdot N(\beta_{1,s} | \mu_{\beta_1}, \sigma_{\beta_1}^2) \cdot N(\gamma_{1,s} | \mu_{\gamma_1}, \sigma_{\gamma_1}^2) \cdot N(\alpha_{2,s} | \mu_{\alpha_2}, \sigma_{\alpha_2}^2) \cdot N(\beta_{2,s} | \mu_{\beta_2}, \sigma_{\beta_2}^2) \cdot N(\gamma_{2,s} | \mu_{\gamma_2}, \sigma_{\gamma_2}^2) \cdot G(\theta_s | \theta_k, \theta_h) \cdot G(\sigma_{t,s} | \sigma_k, \sigma_\theta) \cdot N(\mu_{\alpha_1} | 0, 1) \cdot G(\sigma_{\alpha_1} | 0.1, 10) \cdot N(\mu_{\beta_1} | 0, 1) \cdot G(\sigma_{\beta_1} | 0.1, 10) \cdot N(\mu_{\gamma_1} | 0, 1) \cdot G(\sigma_{\gamma_1} | 0.1, 10) \cdot N(\mu_{\alpha_2} | 0, 1) \cdot G(\sigma_{\alpha_2} | 0.1, 10) \cdot N(\mu_{\beta_2} | 0, 1) \cdot G(\sigma_{\beta_2} | 0.1, 10) \cdot N(\mu_{\gamma_2} | 0, 1) \cdot G(\sigma_{\gamma_2} | 0.1, 10) \cdot G(\theta_k | 1, 10) \cdot G(\theta_h | 1, 10) \cdot G(\sigma_k | 1, 10) \cdot G(\sigma_\theta | 1, 10) \quad (8)$$

모든 매개변수에 대한 다중 적분을 직접적으로 추정하는 것은 불가능하며 본 연구에서는 Bayesian Markov Chain Monte Carlo (MCMC) 방법을 도입하여 깃스표본법을 이용하여 각 매개변수의 사후분포를 추정하였다. Bayesian MCMC 기법은 사후분포를 추정하는 데 있어서 다변량에 대한 복잡한 적분을 위해서 적용되는 수치해석 기법으로 매개변수를 추정하는 수단으로 활용되고 있다(Kwon *et al.*, 2008). 주어진 다변량 확률분포가 복잡하여 이를 따르는 Independent And Identically Distributed (IID) 난수를 얻을 수 없는 경우에 사용가능한 기법으로서 IID 난수 대신 Markov Chain 난수를 추출하여 사용한다. Markov Chain을 통해 다변량 확률분포를 따르는 난수를 재현하기 위해서는 반복 시행을 통해 분포에 수렴시키는 과정이 필요하다. MCMC 기법의 대표적인 방법으로는 메트로폴리스-헤스팅 알고리즘과 깃스표본법을 있으며, 본 연구에서는 깃스표본법을 이용하였다. 깃스표본법에 대한 자세한 내용은 기존 연구문헌들을 참조할 수 있다(Kwon *et al.*, 2008; Lee *et al.*, 2010).

3. 결 과

3.1 최적 모형의 선정

본 연구에서는 기상인자인 월강수량의 유무로 모형을 구분지어 비교하였다. 모형을 적합하는 과정에서 매개변수가 추가되면 종속변수에 대한 설명력은 개선되지만, 과적합(overfitting)에 대한 문제가 발생한다. 이에 해결하기 위해 매개변수 개수에 벌점(penalty)을 적용하는 방법을 도입하였으며, 모형 선정 기준은 대표적으로 AIC, BIC, DIC 등과 같은 지표들이 존재한다. Akaike (1974)이 제안한 Akaike Information Criterion (AIC)는 확률분포 간의 차이를 의미하는 Kullback-Leibler 거리에 의해 정의되며 모형의 대수우도함수와 모형의 복잡한 정도를 나타내는 매개변수의 개수로 표현된다. AIC는 대수우도와 매개변수 개수를 기반으로 산정되며, 큰 우도값과 적은 개수의 매개변수로 이루어져 지표가 작게 산정된 모형이 최적모형이며, 통계적으로 우수한 모형을 의미한다. Eq. (9)은 우도를 나타내며 Eq. (10)은 AIC 지표를 나타낸다.

$$L = \ln \left(\prod_{i=1}^n f(x_i | \theta) \right) \quad (9)$$

$$AIC = -2 \ln(L) + 2k \quad (10)$$

Eq. (9)에서 n 은 자료의 수를 의미하며, f 는 매개변수 θ 를 가지는 확률밀도함수를 나타낸다. x 는 벡터형태의 자료를 의미하고, Eq. (10)에서 k 는 모형의 매개변수 개수를 나타낸다.

BIC은 Bayesian 사후확률의 형식으로 정의되는 정보기준이며, AIC와 마찬가지로 대수우도와 매개변수 개수로 표현된다. AIC와 다르게 고정된 값이 아닌 표본의 크기에 따라서 벌점을 적용하여 복잡한 모형에 대해서 AIC보다 벌점이 크게 산정된다. AIC와 마찬가지로 지표가 작게 산정된 모형이 최적 모형을 의미한다. Eq. (11)은 BIC 지표를 나타낸다.

$$BIC = -2 \ln(L) + k \ln(n) \quad (11)$$

Eq. (11)에서 k 는 모형의 매개변수 개수, n 은 자료의 개수를 나타낸다.

Spiegelhalter *et al.* (2002)는 Deviance Information Criterion (DIC)을 제시하였으며, 이는 계층적 모형에 대한 AIC와 BIC 기준을 일반화한 형태이다. 특히, MCMC 기법에 의해서 사후분포가 얻어지는 Bayesian 모형에서 적합한 모형 선택 시 유용하게 사용된다. DIC는 편차(deviance) $D(\theta) = -2 \log(p(y|\theta))$

의 평균값 $\overline{D(\theta)}$ 와 매개변수 개수를 의미하는 $p_D = \overline{D(\theta)} - D(\bar{\theta})$ 로 결합된다. Eq. (12)은 DIC 지표를 나타내며, 앞서 소개한 AIC, BIC와 동일하게 작게 산정된 모형이 통계적으로 최적 모형을 의미한다.

$$DIC = p_D + \overline{D(\theta)} \quad (12)$$

본 연구에서는 우도와 매개변수 개수 기반으로 산정되는 AIC와 BIC을 기준으로 최적모형을 선정하였다. 독립변수가 추가됨으로써 Two-Stage 모형 특성상 층별로 임계점을 기준으로 두 개의 회귀계수가 추가되었으며, 그에 따라 매개변수의 개수가 증가하였다. 모형에서 임계점의 역할을 알아보고자 임계점을 갖지 않는 One-Stage모형의 AIC, BIC를 Table 2에 비교하여 나타내었다. 전체적으로 Two-Stage 모형의 지표값이 낮게 추정되고 있으며, 기상인자를 추가한 결과가 지표값의 증가로 나타내고 있다. 이는 추가된 독립변수가 나타내는 설명력보다 매개변수 개수에 대한 벌점이 크게 작용한 결과를 의미한다. 본 연구에서는 계절에 따라 크게 변동하는 염분자료의 양상을 효과적으로 재현하기 위하여 임계점을 가지는 모형 활용이 통계적으로 유리한 것을 확인할 수 있다.

구체적으로 값을 비교해보면, One-Stage을 이용한 최적 모형의 BIC값은 483.7를 나타내고 있으나 Two-Stage을 이용한 최적 모형의 경우 BIC값이 366.7으로 추정되는 등 상대적으로 모형의 적합성 측면에서 매우 우수함을 확인할 수 있다. 추가적인 모형 적합성을 판단하기 위하여 AIC값을 산정하였으며, 그 결과 BIC 결과값과 동일하게 나타나 Two-Stage 모형의 적합성을 다시 한번 확인하였다. One-Stage모형은 Two-Stage 모형에 비해 상대적으로 매개변수의 수는 적지만, 모형의 설명력을 나타내는 우도가 작기 때문에 나타나는 현상이라 판단할 수 있다. Fig. 2에서 볼 수 있듯이 염분자료는 큰 변동폭을 갖기에 자료를 분리하지 않고 모형을 구축한다면, 재현성이 떨어지고 극치사상을 재현하는 것이 어려울 수 있다. 따라서, Two-Stage 모형이 상대적으로 우수한 성능을 가지는 것을 확

Table 2. A comparison of model performance with BIC and AIC values over four different models

Model		Salinity Simulation with Different Predictors	
Independent variable		Y_{t-1}, X_t	Y_{t-1}, X_t, R_{t-1}
BIC value	Two-Stage	366.7	389.6
	One-Stage	483.7	495.8
AIC value	Two-Stage	308.9	317.4
	One-Stage	454.8	459.7

인할 수 있으며, 최종적으로 이전기간의 염분 (Y_{t-1})과 상시 모니터링 지점의 염분(X_t)으로 구성된 모형을 최적모형으로 선정하였다.

3.2 염분모의모형 구축 및 적합

본 연구에서는 앞에서 언급했듯이 임계점으로 구분된 Two-Stage ARX 모형을 적용하였으며, 계층적 Bayesian 기법을 사용하여 매개변수의 사후분포를 추정하였다. 앞서 언급하였듯이 월평균 염분자료를 시계열 형태로 구축한 후 전처리 과정인 Log-scale 변환과 정규화를 수행하였다. Table 3에 추정된 매개변수의 평균과 표준편차를 제시하였으며, 회귀계수를 나타내는 $\alpha_{1,s}, \beta_{1,s}, \gamma_{1,s}, \alpha_{2,s}, \beta_{2,s}, \gamma_{2,s}, \theta_s, \sigma_{t,s}$ 는 층(s)에 따라 추정되었다.

Table 3. Estimated parameters of salinity simulation model for the entire period

Parameter	Mean			Standard Deviation		
	Surface	Middle	Bottom	Surface	Middle	Bottom
$\alpha_{1,s}$	0.17	0.14	0.18	0.11	0.12	0.10
$\alpha_{2,s}$	0.03	0.09	0.11	0.09	0.08	0.08
$\beta_{1,s}$	0.14	0.13	-0.01	0.10	0.10	0.09
$\beta_{1,s}$	0.16	0.16	0.20	0.07	0.08	0.08
$\gamma_{1,s}$	0.67	0.71	0.66	0.09	0.10	0.09
$\gamma_{2,s}$	-0.78	-0.71	-0.76	0.09	0.09	0.09
θ_s	0.01	0.04	0.02	0.04	0.05	0.03
$\sigma_{t,s}$	4.21	3.76	3.75	0.61	0.53	0.53
Parameter	Hyper-parameter Common Mean		Hyper-parameter Common Standard Deviation			
μ_{α_1}	0.15		0.30			
σ_{α_1}	0.07		0.28			
μ_{α_2}	0.08		0.28			
σ_{α_2}	0.16		0.28			
μ_{β_1}	0.63		0.30			
σ_{β_1}	-0.69		0.30			
μ_{β_2}	10.22		9.51			
σ_{β_2}	10.83		9.98			
μ_{γ_1}	10.36		9.56			
σ_{γ_1}	10.58		9.67			
μ_{γ_2}	10.40		9.80			
σ_{γ_2}	10.57		10.17			
θ_k	22.03		12.85			
θ_h	5.74		3.40			
σ_k	0.55		0.72			
σ_θ	16.10		12.63			

모형의 적합성을 평가하기 위해서 본 절에서는 전체기간 2011년-2017년에 대해서 통계지표를 이용하여 모의결과를 비교하였다. 염분모의 모형을 통해 모의된 값과 동진 지점에서 관측된 값 사이에 적합성을 정량적으로 평가하고자 층별(표층, 중층, 저층)로 나누어 상관계수(Correlation Coefficient, CC), 일치계수(Index of Agreement, IoA), 평균제곱근오차(Root Mean Square Error, RMSE), 평균절대오차(Mean Absolute Error, MAE) 등을 활용하였다. 상관계수는 관측자료와 모형결과간의 선형관계를 나타내는 무차원으로 1에 가까울수록 모의능력이 우수함을 나타낸다. 일치계수는 평균과 분산사이의 차이를 추정할 때 상관계수의 근거한 통계치의 민감성을 보완하고자 Willmott (1981)에 의해 개발되었다. 평균 제곱근 오차는 실제 값과 모형결과값의 차이를 나타내는 척도이며 0에 가까울수록 모형의 우수함을 나타낸다. 평균절대오차는 관측값과 모형값 오차의 절댓값 평균으로 산정되며 0에 가까울수록 우수함을 의미한다. Eq. (13)은 위의 통계적 평가 지표를 나타내며 Table 4에 층별 통계적 분석 결과를 나

타내었다.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (S_i - O_i)^2} \tag{13a}$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |S_i - O_i| \tag{13b}$$

$$IoA = 1 - \frac{\sum_{i=1}^N |O_i - S_i|^p}{\sum_{i=1}^N (|S_i - \bar{O}| + |O_i - \bar{O}|)^p} \tag{13c}$$

$$CC = \frac{\sum_{i=1}^N (S_i - \bar{S})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^N (S_i - \bar{S})^2 \sum_{i=1}^N (O_i - \bar{O})^2}} \tag{13d}$$

Table 4. Statistical evaluation of salinity simulation model for the entire period

Layer	CC	IoA	MAE (unit : psu)	RMSE (unit : psu)
Surface	0.7935	0.8812	1.0279	2.1682
Middle	0.7932	0.8847	1.2696	2.5470
Bottom	0.8382	0.9049	1.3002	2.4510

Table 4의 결과를 살펴보면, 모든 층에서 상관계수는 약 0.8 정도로 높은 상관성을 나타내었으며 일치계수는 약 0.88로 산정되어 높은 일치도를 보였다. 전반적인 결과를 비교해보면, 염분모의 모형으로부터 모의된 값은 관측자료와 높은 상관성을 보이며 상관계수, 일치계수, 평균제곱근 오차, 평균절대오차 등의 통계적 지표로부터 우수한 모의능력을 나타내는 것을 확인할 수 있다. Fig. 4은 최적모형으로 선정된 Two-Stage ARX

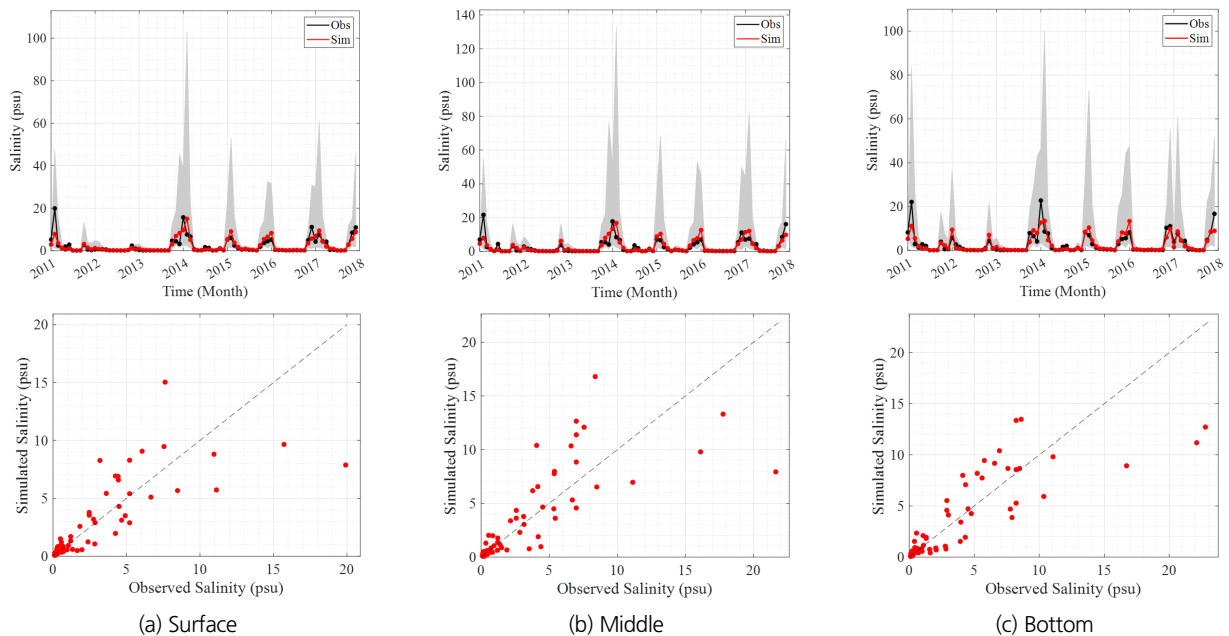


Fig. 4. A comparison of model performance and scatter plots on salinity simulation for three layers for the entire period

모형의 층별 모의된 중앙값(median)과 Fig. 2에 나타난 관측값을 도시하였으며, 산점도 또한 Fig. 4 우측에 나타내었다.

3.3 염분모의 검정 및 검증

2.3절에서 서술한 계층적 Bayesian 추론을 통해 모형의 매개변수를 확률분포형태로 산정하였으며, MCMC기법의 깃스표본법을 이용하여 다변량 확률분포를 따르는 표본을 생산하였다. 매개변수의 수렴성을 확인하기 위해 Bayesian 모형 내에서 3개의 Chain을 독립적으로 시행하였으며, 7,000번 모의 발생 중 5,000번은 제거(burn-in)하고, 나머지 2,000개의 표본을 확충하여 각 매개변수별 사후분포를 추정하였다.

모의의 적합성을 평가하기 위하여 모형 검증(validation)을 수행하였으며, 이를 위해 시계열 자료를 분리하였다. 분리한 자료를 사용해 모형을 보정(calibration)하였으며, 보정을 통해 추정된 매개변수를 분리된 일부분 자료에 대해서 검증하였다. 즉, 일부분의 표본을 학습에 사용하였으며, 학습에 사용되지 않은 표본을 통해 모형 검증을 하였다. 본 연구에서는 전처리

를 수행한 염분자료를 2011-2015년(5년)과 2016-2017년(2년)으로 구분하여 5년 기간의 자료를 활용해 매개변수를 추정하였고 나머지 2년 기간의 자료를 활용해 모형을 검증하였다.

Fig. 5는 염분모의모형을 통해 추정된 중앙값과 염분 관측값을 층별로 구분하여 비교한 그래프이다. 겨울철에 높아지는 염분패턴을 효과적으로 재현하고 있으며 낮은 염분에 대한 높은 재현성과 극치 염분의 범위까지 효율적으로 재현하는 것을 확인할 수 있다. 또한, 매개변수의 사후분포를 활용하여 불확실성 범위를 제시하였다. Fig. 5에서 제시된 Uncertainty는 Bayesian 모형으로부터 추정되는 Predictive Posterior 분포로부터 추정되는 Credible Interval로 기존 Confidence Interval(CI)과는 다르게 해석된다. 기존 Confidence Interval은 대상 자료를 정규분포로 가정하여 분석하였기에 좌우 대칭의 형태를 가지며, 본 연구에서 고려된 Bayesian 방법에서는 Lognormal 분포로 고려하였기에 독립 및 종속변수의 크기에 따라 불확실성이 다르게 추정되는 이분산성(heteroscedasticity)을 가진다. 즉, 자료가 크기가 커지면 분산이 커지는 특성을 가지며, 이러한 특성을 모델링 과정에서 고려하여 계산된 Predictive Posterior 분포를 그림으로 표시하였다. 본 연구에서 제시하는 Bayesian 모형은 기본적으로 매개변수의 불확실성을 추정하기 위한 목적으로 적용된다. 모형 구성 시 관측오차를 포함하여, 매개변수 및 모형의 불확실성 등 3개로 분리할 수 있다. 이에 본 연구에서는 매개변수의 불확실성을 정량화하였으며, 매개변수의 불확실성을 제외한 불확실성은 모형 및 관측 불확실성이라고 간주할 수 있다. 다만 Fig. 5의 결과를 판단하

Table 5. Statistical evaluation of salinity simulation model for validation period

Layer	CC	IoA	MAE (unit : psu)	RMSE (unit : psu)
Surface	0.8764	0.9329	0.9846	1.7235
Middle	0.8291	0.8967	1.2772	2.4271
Bottom	0.8721	0.9260	1.1521	2.2656

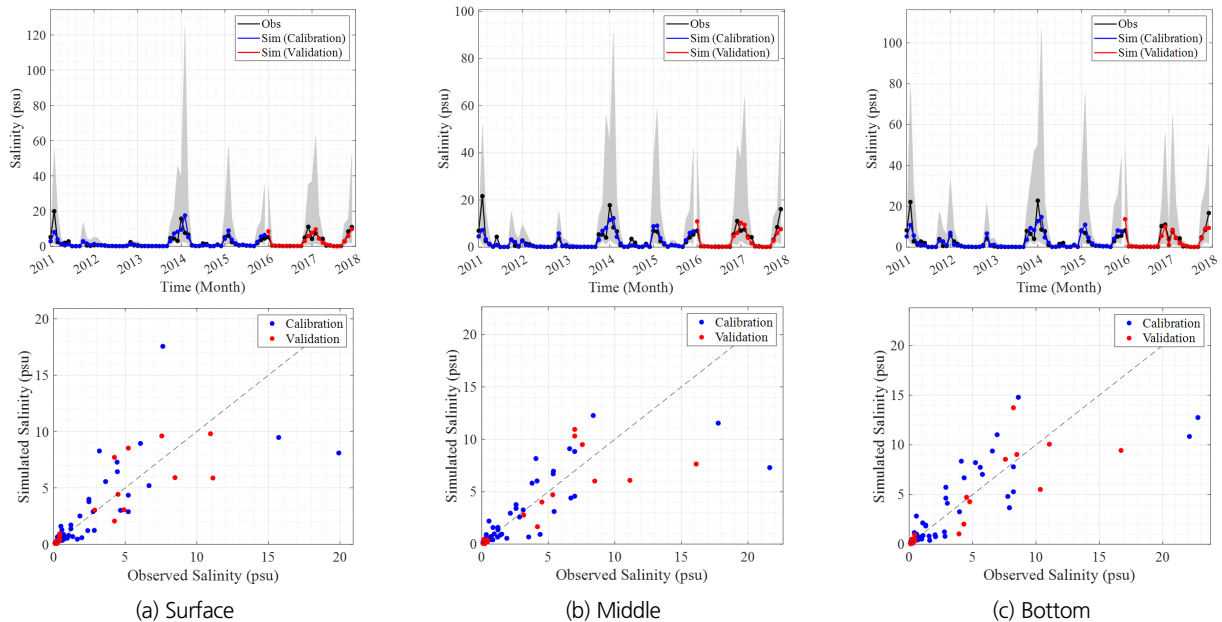


Fig. 5. A comparison of model performance and scatter plots on salinity simulation results from calibration and validation phases for three layers

면 대부분의 관측값이 제시된 불확실성 구간에 위치하여 관측 오차는 매개변수 오차에 비해 매우 작음을 인지할 수 있다. Fig. 5의 우측에는 앞서 비교한 염분모의모형의 모의된 중앙값과 염분 관측값의 산점도를 검정과 검증으로 구분하여 나타내었으며, 검증 부분에 대한 상관계수를 포함한 통계적 평가 지표를 Table 5에 정리하였다. 앞서 모형 구축결과와 유산하게 검정 및 보정으로 구분한 모형 적합성 결과도 매우 우수함을 확인할 수 있다.

4. 결론 및 고찰

새만금호내로 유입되는 만경·동진 수역에 양수장을 운영하여 농업용수를 공급할 예정이지만, 강수량이 적은 겨울철에는 염분농도가 큰 폭으로 상승하여 농업용수로서의 이용할 수 없어 양수장운영에 어려움을 발생시킨다. 배수갑문을 통해 유통되는 해수가 양수장까지 미치는 영향을 사전에 대비한다면, 그에 따른 양수장 운영방안을 제시할 수 있을 것으로 판단된다. 따라서 본 연구에서는 예측된 염분 시계열자료를 활용하여 추계학적 염분모의모형을 제시하였으며, 자료가 층별로 구분되어 있어 계층적 Bayesian 기법을 적용하여 매개변수를 확률분포로 추정하였다. 모형의 정량적 평가를 위해 통계적 평가를 수행하였으며, 그 결과 우수한 것으로 판단하였다. 본 연구를 통하여 도출된 결론은 다음과 같다.

- 1) 염분의 자료 특성상 자기상관성이 높은 자료이기에 자기상관회귀모형을 기반으로 하는 ARX모형을 사용하였으며, 계절적인 변동성이 큰 폭으로 변화하는 양상을 보이며 이를 재현하기 위해 임계점을 설정하여 두 가지로 구분하는 Two-Stage ARX모형을 제시하였다. 상대적으로 짧은 기간의 자료를 사용해서 극치사상에 대한 불확실성구간이 크게 산정되었지만, 하나의 모형을 이용한 결과에 비하여 우수한 재현능력을 확인할 수 있었다.
- 2) 기상인자인 월강수량을 고려한 모형을 추가적으로 구축하였지만, 모형의 지표를 나타내는 BIC값이 크게 산정되어 독립변수에 월강수량 포함하지 않는 모형이 통계적 측면에서 우수하다고 판단하였다. 이는 결과에서 보았듯이 기존의 독립변수들이 종속변수에 대한 설명력이 크기에 기상인자가 추가되어도 큰 변화가 없으며, 매개변수가 늘어나는 효과가 상대적으로 크게 작용하였다. 따라서 기상인자를 포함하지 않는 모형이 우수하다고 판단하였다.

본 연구를 수행하면서 짧은 기간의 자료를 가지고 모형을

구축하는 것이 한계점으로 나타났다. 입력자료의 부족이라는 문제를 해결하기 위해 생산된 표본을 추가적으로 필터링한 후 활용하는 방안의 연구가 필요하다. 또한, 새만금내 동진양수장 수역 염분에 염분모의 모형을 적용하여 모의 확인을 수행한 결과를 토대로 추후 염분 예측모형을 개발하면, 양수장 운영을 체계적이고 합리적으로 운영방안을 제시할 수 있다고 판단된다. 추가적으로 자료의 시공간적 규모의 차이를 극복하기 위한 방법으로 수치모형, 모니터링 정보 등을 통합하여 일단위 염분 시공간적 상세화 모형으로 확장이 가능하며, 이를 활용한 염분 농도 예측 및 전망이 가능할 것으로 판단된다.

감사의 글

본 연구는 농림축산식품부의 재원으로 한국농어촌공사 농어촌연구원의 새만금 조사연구 사업의 지원을 받아 연구되었습니다.

References

- Akaike, H. (1974). "A new look at the statistical model identification." *IEEE transactions on automatic control*, Vol. 19, No. 6, pp. 716-723.
- Bauder, T.A., Waskom, R.M., Sutherland, P.L., Davis, J.G., Follett, R.H., and Soltanpour, P.N. (2011). *Irrigation water quality criteria*. Colorado State University Extension, No. 0.506.
- Cho, H.Y., Cho, B.J., and Kim, S.J. (2008). "Estimation on the regions of freshwater influence using the seasonal salinity data in Asan Bay coastal zone." *Journal of Korean Society of Coastal and Ocean Engineers*, Vol. 20, No. 2, pp. 219-231.
- Cho, Y.M., Ryu, D.C., Lee, K.S., and Park, J.R. (2018). "Strategies to defense of intake raw water according to open of Nakdong River estuary gate" *Journal of Korean Society of Water Science and Technology*, Vol. 26, No. 1, pp. 47-54.
- Gelman, A. (2006). "Prior distributions for variance parameters in hierarchical models." *Bayesian analysis*, ISBA, Vol. 1, No. 3, pp. 515-534.
- Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. (2004). "Bayesian data analysis." Chapman and Hall/CRC. Inc., N.Y., U.S.
- Jeong, Y.H., and Yang, J.S. (2015). "The long-term variations of water qualities in the Saemangeum salt-water lake after the sea-dike construction." *Journal of the Korean Society for Marine Environment and Energy*, Vol. 18, No. 2, pp. 51-63.
- Kwon, H.H., Brown, C., and Lall, U. (2008). "Climate informed flood frequency analysis and prediction in Montana using hierarchical Bayesian modeling." *Geophysical Research Letters*,

- American Geophysical Union*, Vol. 35, No. 5, L05404.
- Kwon, H.H., Kim, J.Y., Kim, O.K., and Lee, J.J. (2013). "A development of regional frequency model based on hierarchical bayesian model." *Journal of Korea Water Resources Association*, Vol. 46, No. 1, pp. 13-24.
- Lee, J.J., Kwon, H.H., and Kim, T.W. (2010). "Concept of trend analysis of hydrologic extreme variables and nonstationary frequency analysis." *Journal of the Korean Society of Civil Engineers*, Vol. 30, No. 4B, pp. 389-397.
- Lettenmaier, D.P., and Wood, E.F. (1993). "Hydrologic forecasting." *Handbook of hydrology*, McGraw-Hill, N. Y., pp. 26.1-26.30.
- Marshall III, F.E., Smith, D., and Nickerson, D. (2004). *Using statistical models to simulate salinity variation and other physical parameters in north Florida Bay*. Cetacean Logic Foundation. Inc. New Smyrna Beach, F.L., U.S.
- Ministry of Agriculture, Food and Rural Affairs (MAFRA) (2017). *Design strategy of demand customized agricultural water supply system in Saemangeum Project*.
- Ministry of Environment (ME) (2011). *Study on the operation of environment-friendly drainage sluice gate to improve the water quality of Saemangeum*.
- Park, J.S. (2002). *Application of ARIMA Model for Han River water quality analysis and forecasting*. Master thesis, Yonsei University.
- Qiu, C., and Wan, Y. (2013). "Time series modeling and prediction of salinity in the Caloosahatchee River Estuary." *Water Resources Research, American Geophysical Union*, Vol. 49, No. 9, pp. 5804-5816.
- Salas, J.D. (1980). *Applied modeling of hydrologic time series*. Water Resources Publication, pp. 6-16.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and Van Der Linde, A. (2002). "Bayesian measures of model complexity and fit." *Journal of the royal statistical society: Series b (statistical methodology)*, Vol. 64, No. 4, pp. 583-639.
- Suh, S.W., and Lee, H.Y. (2008). "Water quality simulations after completion of Saemangeum reservoir construction." *Journal of the Korean Society of Civil Engineers*, Vol. 28, No. 1B, pp. 79-93.
- Wang, Y., Liu, Z., Gao, H., Ju, L., and Guo, X. (2011). "Response of salinity distribution around the Yellow River mouth to abrupt changes in river discharge." *Continental Shelf Research, Elsevier*, Vol. 31, No. 6, pp. 685-694.
- Willmott, C.J. (1981). "On the validation of models." *Physical geography, Taylor and Francis Group*, Vol. 2, No. 2, pp. 184-194.