



Prediction model for electric power consumption of seawater desalination based on machine learning by seawater quality change in future

Shim, Kyudae^a · Ko, Young-Hee^{b*}

^aLeader, Environment Solution Team, GS Engineering & Construction, Seoul, Korea

^bAssociate Professor, AI Big Data MBA, Seoul School of Integrated Sciences & Technologies, Seoul, Korea

Paper number: 21-067

Received: 23 August 2021; Revised: 23 September 2021; Accepted: 27 September 2021

Abstract

The electricity cost of a desalination facility was also predicted and reviewed, which allowed the proposed model to be incorporated into the future design of such facilities. Input data from 2003 to 2014 of the Korea Hydrographic and Oceanographic Agency (KHOA) were used, and the structure of the model was determined using the trial and error method to analyze as well as hyperparameters such as salinity and seawater temperature. The future seawater quality was estimated by optimizing the prediction model based on machine learning. Results indicated that the seawater temperature would be similar to the existing pattern, and salinity showed a gradual decrease in the maximum value from the past measurement data. Therefore, it was reviewed that the electricity cost for seawater desalination decreased by approximately 0.80% and a process configuration was determined to be necessary. This study aimed at establishing a machine-learning-based prediction model to predict future water quality changes, reviewed the impact on the scale of seawater desalination facilities, and suggested alternatives.

Keywords: Machine learning, Seawater temperature, Salinity, Prediction model, Electric power

장래 해수수질 변화에 따른 머신러닝 기반 해수담수 전력비 예측 모형 개발

심규대^a · 고영희^{b*}

^aGS건설 환경솔루션팀 책임, ^b서울과학기술대학교 AI빅데이터 MBA 부교수

요 지

본 연구는 머신러닝 기반의 분석으로 해수담수화(Desalination) 시설의 전력비 예측모델의 가능성을 검토하였다. 해수담수화 주요 공정인 역삼투(Seawater Reverse Osmosis) 시설의 전력비 예측 모델을 개발하고, 전력비 산정에 영향을 미치는 인자를 분석하였으며, 해수 수질 중에서 선정된 수온 및 염분도 측정자료를 활용하여 검토하였다. 국립해양조사원(Korea Hydrographic and Oceanographic Agency, KHOA)의 2003년부터 2014년까지의 자료를 이용하였으며, 모형의 구조는 시행오차법(Trial & Error)으로 하이퍼파라미터를 최적화하여 머신러닝 기반의 예측 모델을 구축하고, 장래 해수 수질을 예측하였다. 해수 수온은 기존 패턴과 유사할 것으로 예측되었고, 염분도는 과거 측정자료 범위 이내로 최대값이 점차 감소되는 경향을 보여 해수담수화의 전력비가 약 0.80% 감소하는 것으로 검토되었다. 본 연구는 머신러닝 기반의 예측 모델을 구축하여 장래 수질 변화 예측하였으며, 해수 수질 변동의 영향 및 대안을 제시했다는데 의의가 있다.

핵심용어: 머신러닝, 해수 수온, 염분도, 예측 모델, 전력비

*Corresponding Author. Tel: +82-70-7012-2724
E-mail: yhko@assist.ac.kr (Y. H. Ko)

1. 서론

우리나라는 세계보건기구(World Health Organization, WHO)에 보고된 물 부족 국가 중 하나로, 최근에는 이상 기후로 인하여 가뭄이 자주 발생하며, 그 규모 및 피해도 증가하고 있다. 여름철에 태풍 및 강우가 집중적으로 발생하는 우리나라 기후 특성상, 대부분의 수자원이 하천을 통해 바다로 방류되므로 여름철을 제외한 기간은 물이 부족한 실정이며, 환경 오염으로 인한 사용할 수 있는 수자원 또한 점차 감소하고 있어, 물 부족으로 새로운 대체 수자원 개발이 논의되고 있다. 다양한 대체 수자원 중 해수담수시설이 해결 방안의 하나로 검토되고 있으며, 우리나라는 삼면이 바다와 접해 있어, 지리학적으로도 해수담수화 추진이 용이하다(Kim and Oh, 2008; K-water, 2019; Lee and Park, 2020).

최근 반도체 및 중화학산업 기술의 발전으로 인한 산업용수 수요는 급격히 증가하고 있으며(K-water, 2019; Kwon *et al.*, 2020), 산업용수가 확보되지 않으면 산업단지 개발이 될 수 없기에 수자원 확보가 최우선적으로 고려되기 때문에, 안정적이고 양질의 수자원을 지속적으로 공급할 수 있는 해수담수화 공법은 물 공급의 좋은 해법이 될 수 있다.

해수담수화 공법은 증발(Thermal Process) 및 막(Membrane Process)으로 구분되며, 기존 해수담수화 연구는 전처리 공법, 해수담수화 공정별 전력비 소모율, 역삼투압 공정 전력비 및 운영 전력비 효율 최적화 등에 대해서 주로 이루어졌다. 전력비(Specific Energy Consumption, SEC)는 가정 또는 산업단지에 물 공급을 위한 생산수의 체적 당 필요한 전력(kWh/m³)을 말한다. 최근 해수담수화 시설 경향은 경제적인 역삼투(Reverse Osmosis, RO) 공법으로 주로 적용되며(Greenlee, 2009), Desalination Market 2016(GWI, 2015)에 의하면 향후 발주되는 프로젝트의 약 90%가 역삼투 공법으로 진행될 것으로 예측하였고, 증발법에서 막공법으로 공정이 변화하면서, 해수담수화에 사용되는 전력은 22 kWh/m³에서 3 kWh/m³까지 낮아져, 약 90% 정도의 절감효과가 있다고 평가했다(Bernat *et al.*, 2010; Badruzzaman *et al.*, 2019). Kurihara and Takeuchi (2013)은 해수담수화 시설의 에너지를 절감하기 위해 저압막 개발, 에너지 회수장치(Energy Recovery Device, ERD) 및 역삼투 공정 시스템의 필요성을 검토하였으며, Fritzmann *et al.* (2007)은 해수담수화 공법별 에너지 비용을 조사하고, 많은 유량을 통과시킬 수 있는 막 및 에너지 회수 장치의 개발로 전력비가 낮아졌다고 분석하였다. Voutchkov (2018)는 해수담수에 소요되는 에너지는 취수공정 5%, 전처리공정 10%, 역삼투공정 70%, 기타 25%로 평가하였으며, 전체 전력비 중 역삼투 공정 전력비 비중이 가장 높은 것으로 연

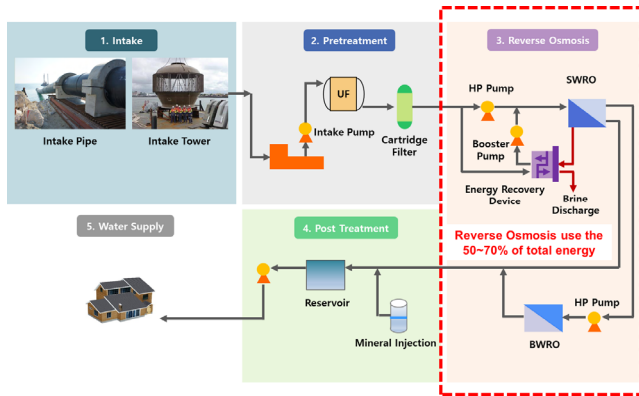
구하였다. 일반적인 해수담수화 소요 전력은 기수 0.3-2.6 kWh/m³, 해수 2.5-4.0 kWh/m³로 나타나고 있으며, 전력비 예측에 주요인자인 해수 수질은 태평양, 대서양 및 지중해는 TDS 35-38 psu, 수온 9-35°C, 중동지역은 TDS 40-45 psu, 수온 16-35°C로 적용되며, 막 제조사의 프로그램을 통해 전력소모량(kWh/m³)을 평가한다(Voutchkov, 2010). 최근에는 인버터펌프 운전으로 에너지를 절감하고 해수담수시설의 운영 최적화를 통한 운영비 절감에 대한 연구가 진행되고 있다(Woo and Kim, 2017; Suh *et al.*, 2015).

기자재 개발 및 다양한 해수담수 관련 공법의 개선으로 해수담수화에 사용되는 전력비는 점차 감소되고 있으며, 운영비를 낮추기 위해 다양한 기술을 도입하고 있다. 그러나 이러한 시설 운영에 필요한 해수담수화 전력비는 낮아지더라도 그 한계가 있으며, 최근에는 기상이변으로 인한 해양수질 조건이 많이 변화하고 있어, 이를 고려한 최적화된 시설물 계획이 필요하다. 특히, 인프라 시설은 그 특성상 한번 시공이 완료되면 최소 20~30년간 지속적인 운영이 되며, 시설 확장 또는 주요 공정 변경이 쉽지 않으므로, 계획 초기단계부터 자세한 검토가 필요하다. 또한 전력비에 가장 큰 영향을 주는 역삼투 공정(약 50~70%)의 운영 전력비 예측은 매우 중요하다. 본 연구는 머신러닝 기법을 활용하여 해양관측소에서 측정된 해수 수질 자료를 활용하여 장래 해수수온 및 염분도를 예측하였고, 모델 구축시 시행-오차법(Trial & Error)을 통해 최적의 신경망 네트워크 구조를 적용하여, 해수의 변동된 값을 결정된 후 관측된 측정 값과 검증을 실시하였다. 머신러닝 알고리즘을 활용한 해수 수질 예측모델 관련 선행연구들은 주로 해양수질 변화에 대해 조사하고, 해양수질 변화에 따른 해수담수화 시설의 영향을 검토한 사례가 없으므로, 본 연구에서는 장래 해수수질 변화에 따른 해수담수 전력비를 산정하고, 이를 고려한 해수담수화 시설의 적정 설계 방법을 제안하고자 한다.

2. 해수담수화 공정 및 선행연구 고찰

2.1 해수담수화 공정 소개

해수담수화 시설은 Fig. 1과 같이, 취수, 전처리, 역삼투, 후처리 및 용수공급 공정으로 이루어진다. 취수공정은 취수시설을 통하여 해수를 취수하여, 펌프 및 배관으로 전처리설비까지 공급하는 시설이다. 전처리 공정은 역삼투압막의 막힘을 방지하기 위하여, 해수 부유물 및 입자성 물질을 제거하고 역삼투압막으로 처리하기 좋도록 전처리하는 시설이다. 역삼투 공정은 펌프, 역삼투막, 에너지회수로 이루어지며, 역삼투막으로 생산수 및 농축수로 분리되어진다. 생산수는 pH와 경도가 낮아,



(a) Process flow diagram of desalination

(b) RO (Reverse Osmosis) facility

Fig. 1. Process flow diagram (desalination) and photo of RO facility

이를 보완하기 위해 미네랄을 주입하여 소비자에 공급되며, 농축수는 해양으로 방류된다. 이 중 역삼투압 공정은 전체 시설의 약 50~70% 전력을 소비하기 때문에, 역삼투압 공정의 전력비를 예측하는 것은 매우 중요하다. Voutchkov (2010)은 역삼투 공정에서 중요한 인자는 수온 및 염분도로 보았으며, 이는 경험적으로 수온 1°C 증가할 때 유량은 3%가 증가되고, 영향은 막의 투과유량은 압력과 수온에 비례하며, 염분 제거는 수온에 반비례하는 것으로 연구되었다. 기초 입력자료는 국립해양조사원(Korea Hydrographic and Oceanographic Agency, KHOA)의 2003년부터 2014년까지 측정된 자료를 활용하였다. 전력비는 물리화학식에 근거한 삼투압 계산식을 이용하여 산정하였고, 장래 수온 및 염분도는 다양한 머신러닝 알고리즘을 적용하여 최적값을 예측할 수 있는 모형을 구축하였다.

2.2 머신러닝 소개

2.2.1 머신러닝의 기본개념

머신러닝은 신경망 알고리즘의 한 분야로, 최근에는 딥러닝이라는 용어로 많이 사용되고 있다. 신경망이라는 이름에서와 같이 뇌의 신경세포 구조를 모방한 알고리즘이며, 이를 바탕으로 Rosenblatt *et al.* (1957)이 1957년 제안된 퍼셉트론(Perceptron) 이론은 딥러닝 및 머신러닝 알고리즘의 기초가 되었다. 하나의 퍼셉트론은 하나의 신경세포, 즉 뉴런(Neuron)을 의미하며 Fig. 2와 같이 나타낼 수 있다. 퍼셉트론 알고리즘은 뉴런 한 개에 X_1, X_2 의 정보가 들어오게 되면 각각의 가중치 W_1, W_2 를 곱한 후 전체를 더한 값이 출력되며, 결과 값은 $X_1 \cdot W_1 + X_2 \cdot W_2$ 가 된다. 그 결과 값은 다음 차례의 뉴런으로 전달되는 과정을 반복하게 되며 처리하는 정보는 단순하지만, 많은 수의 뉴런이 얹혀서 복잡한 정보를 가공할 수 있게 된다.

신경망은 Fig. 3과 같이, 입력 계층(Input Layer), 은닉 계층

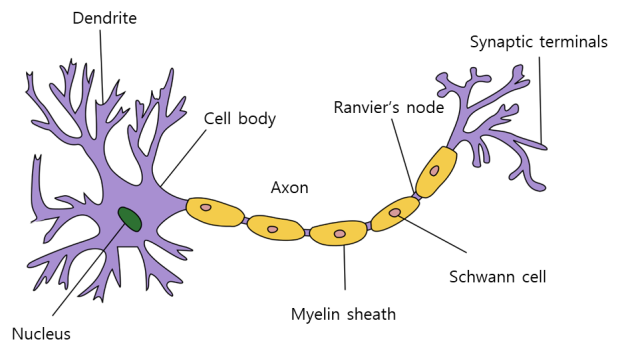


Fig. 2. Biological neuron model (Wikipedia, 2021)

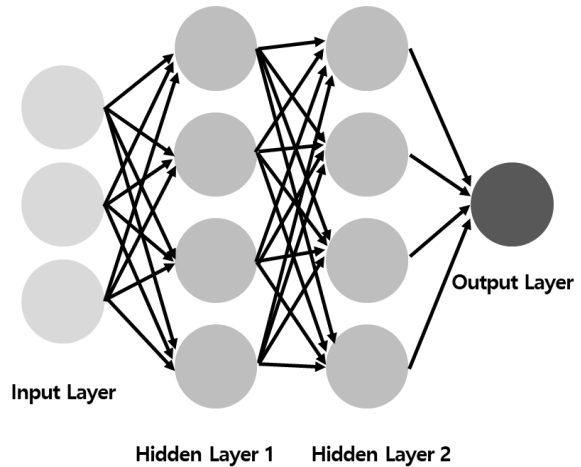


Fig. 3. Basic structure of neural network

(Hidden Layer) 및 출력 계층(Output Layer)으로 나뉘며, 일반적으로 은닉 계층이 한 개일 때 얕은 신경망(Shallow Neural Network), 두 개 이상은 깊은 신경망(Deep Neural Network)이라고 부른다.

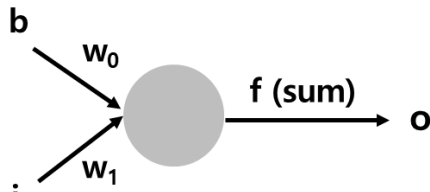


Fig. 4. Diagram of the machine learning process

2.2.2 머신러닝 처리과정

앞에서 언급된 머신러닝 알고리즘을 간단하게 설명하면, 뉴런에 정보가 입력된 후, 가중치에 의해 결과가 조절되어 출력되며, 간소화된 모식도는 Fig. 4와 같다. 뉴런은 노드(Node)로 표기하며, 입력된 값에 가중치 및 활성화 함수를 고려한 출력값을 다음 노드로 전달한다. 가중치 및 바이어스는 입력값을 조정하며, 가중치는 Eq. (1)의 기울기, 바이어스는 상수 역할을 하게 되며, 활성화 함수는 다음 단계로 전달되는 과정을 결정한다. 즉, 노드는 신경세포인 뉴런에서 자극이 역치를 넘을 때 반응하는 기작을 모방하였다(Tariq, 2016).

$$o = f(b \times w_0 + i_1 \times w_1) \tag{1}$$

- o : 출력값(Output)
- f : 활성화함수(Activation function)
- b : 바이어스(Bias)
- w : 가중치(Weight)
- I : 입력(Input)

학습과정에서 각각의 가중치는 관측치와 예측치의 오차를 최소화하도록 가중치를 최적화한다(Cho *et al.*, 2011; Oh, 2008). 학습할 때에는 오차 계산 후 전 단계보다 오차가 큰 경우 학습을 종료하며, 이 과정을 여러 번 반복해 최적화된 가중치를 선정한다(Oh, 2008; Mitchell, 1997). 학습 전에 신경망의 네트워크 구조를 먼저 검토해야 하며, 같은 자료를 적용해도 구조를 어떻게 설정하느냐에 따라 모델의 성능이 달라진다. 이를 조정(Regularization)이라 부르며, 아직은 정형화된 방법이 없기에 시행-오차법(Trial & Error)을 사용한다. 입력 자료 종류, 은닉층 개수, 은닉층의 노드 개수 및 활성화 함수 종류를 정해야 하며, 복잡한 구조와 고급 알고리즘이 항상 좋은 결과가 나타나지 않는다(Oh, 2008). 즉 네트워크를 결정하는 변수가 많으면 학습은 잘 되지만, 일반화 성능이 낮아져 결과 예측력이 떨어지며, 변수가 적으면 적절한 학습의 진행이 어려울 수 있다(Duda *et al.*, 2001).

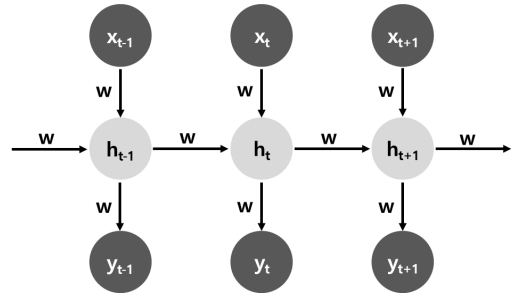


Fig. 5. Basic structure of RNN

2.2.3 RNN (Recurrent Neural Network)

RNN은 순환신경망이라고 하며, 연속된 관측치가 서로 연관성이 있을 때 사용하는 알고리즘이다. 문장 구조 및 시계열 분석에서 많이 사용되며, 특정 시점의 상태를 저장 후, 각 시점의 입/출력에 의해 변화가 발생한다. RNN 기본구조는 Fig. 5와 같고, 이는 Eq. (2)와 같이 나타낼 수 있다.

$$h_t = w_0 + w_1 \cdot h_{t-1} + w_2 \cdot x_t \tag{2}$$

특정 시점의 상태(ht)가 존재하면, 현재 상태(ht)는 직전의 상태(ht-1)과 현재 입력 데이터(xt)에 영향을 받고, 현재 상태(ht)에 따라 출력(yt)이 결정된다. 즉, RNN 알고리즘은 상태를 가정하고, 이에 대한 순차적인 의존성을 나타내며, 다른 머신러닝 알고리즘과 가장 큰 차이점이다. RNN은 긴 시간의 자료를 기억하기 어려우며, 오래 학습할 경우에는 학습 시간이 길어지는 단점이 있다. 이는 기울기 소실(Gradient Vanishing) 문제로서 RNN이 tanh 함수로 학습하기 때문에 단기간 학습 시에는 문제가 없지만, 장기간 학습 시에는 기울기가 0에 가까워져 학습 시간이 느려지고, 효율적으로 학습하지 못하여 학습이 마무리되지 않고 수렴하는 문제가 생긴다.

2.2.4 RNN-LSTM (Recurrent Neural Network-Long Short Term Memory)

LSTM (Long Short Term Memory)은 RNN의 단점을 해결하기 위해 등장한 알고리즘이며, RNN이 계층이 깊어질 경우, 과거 자료를 학습하지 못하고 최근 자료만 기억하여 학습이 잘 안되는 기울기 소실 문제를 해결하였다. RNN 알고리즘과 가장 큰 차이점은 전 단계의 정보를 계속 전달할지 판단하는 구조가 있는 점이다.

LSTM 알고리즘은 Fig. 6과 같이 나타낼 수 있으며, RNN의 기본 구조(Fig. 6)와 달리 복잡한 구조이며, 입력과 현재 상태를 활용해 내부 정보를 업데이트할지 결정한다. 새로운 자료가 유용할 경우는 업데이트가 적용되고, 별로 유용하지 않

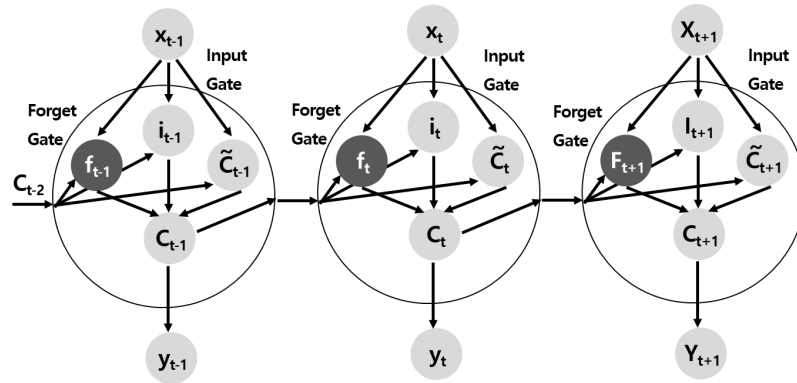


Fig. 6. Basic structure of LSTM

면 지난 단계의 정보를 많이 활용하여 업데이트를 최소화한다. 가장 왼쪽 구조는 망각 게이트로 지금까지 기억한 상태 및 현재 입력값에 계수를 곱해 상태를 전달할지 결정하고, 다음은 입력 게이트로 상태 및 현재 입력값으로 얼마나 전달할지 정하여, 다음 상태를 RNN과 유사하게 추정하게 된다. 이렇게 망각 게이트, 입력 게이트 및 RNN 전과정을 통하여 계산이 되면, 어떤 값을 얼마나 업데이트 할지 정하게 되고 출력은 현재 상태(C_t)와 입력을 적용해서 결정한다.

2.2.5 머신러닝기반 예측모델 알고리즘 및 관련 선행연구

현재 다양한 분야에서 머신러닝과 관련된 연구가 진행되고 있으며, 특히 수질 예측과 관련되어 많은 검토가 되고 있다. Park *et al.* (2020)은 LSTM 알고리즘을 적용한 하천 고탁도 예측(> 30 NTU)을 통한 상수도 공급 최적화를 연구하였다. Jun *et al.* (2020)은 낙동강 하류의 유량 및 수질(Chlorophyll-a, BOD, COD, SS) 예측을 하고 신경망 모델이 BRNN 및 SVM 보다 높은 정확도를 가지는 것으로 검토하였으며, 이와 유사한 연구로 낙동강 하류지역 Chlorophyll-a의 수질 예측을 부스팅(Boosting), 디시전트리(Decision Tree), 랜덤포레스트(Random Forest) 등의 알고리즘을 적용하여 비교하였다(Lee and Kim, 2021).

기존 해양 수질 모델링의 경우, 복잡한 인자를 바탕으로 개발되어 정확도가 낮고 구축 비용 및 시간이 많이 소요되었지만, 머신러닝 기반 모델은 수온 및 수질에 대한 물리적인 이해가 필요하지 않기 때문에 최근 많이 활용되고 있다. 해수 수온 및 수질을 추정하기 위해 머신러닝 기반의 모델이 다양하게 연구되고 있고, Jeon *et al.* (2020)는 머신러닝을 적용한 수질평가 지수 예측연구를 수행하였으며, 랜덤포레스트 및 서포트벡터머신(Support Vector Machine) 알고리즘이 다른 머신러닝 알고리즘보다 우수한 성능을 나타내는 것으로 연구하였다. 기존 SVM과 전통적인 회귀 알고리즘은 장기 시계열 패턴을 예

측하는데 상대적으로 낮은 성능을 보였다. 반면, 시계열 자료를 모의하는데 특화된 LSTM을 사용한 선행연구들에서 기존의 통계적 방법과 비교해서 상대적으로 우수한 성능을 보였다(Zhang *et al.*, 2017). 또한 합성곱신경망(Convolutional Neural Network)을 LSTM과 단순 결합한 모델인 Fully Connected LSTM (FC-LSTM)이 해수면 온도 예측연구에서 SVM보다 상대적으로 우수한 해수면온도 예측 성능을 보여주었다(Yang *et al.*, 2017). Kim *et al.* (2014)은 랜덤포레스트, 큐비스트 회귀(Cubist Regression), 서포트벡터 회귀(Support Vector Regression)를 적용하여 위성 데이터 자료로 서해안의 수질 농도를 추정하였다.

최근 연구는 위성영상 자료인 SST (Sea Surface Temperature) 데이터를 이용한 고수온 탐지 예측을 LSTM 알고리즘을 적용하여 분석하였으며(Choi *et al.*, 2020; Jung *et al.*, 2020), 기상청 자료를 활용하여 해양수질 및 수온을 서포트벡터머신(Support Vector Machine), 랜덤포레스트(Random Forest) 및 LSTM 알고리즘을 적용하여 예측하는 연구를 수행하였다(Jeong and Park, 2020). 다만 이들 연구는 기간이 30분에서 5일 이내로 비교적 단기간 예측 또는 고수온 탐지만 판별한 점은 연구의 한계로 나타났다.

3. 모델 구축

3.1 대상지역의 선정

대상지역(Fig. 7)인 대산항은 서해 태안반도에 위치하며, 주변에는 예산군, 당진시, 태안군 및 홍성군이 인접하고 있다. 석유화학 및 자동차산업이 주요 산업으로, 1991년 무역항으로 지정되어 LG화학, 한화토탈, 현대오일뱅크, 현대OCI 등의 산업체가 있어 산업용수의 공급이 많이 요구되는 지역이다. 따라서 본 연구의 연구대상 지역으로 적합성을 갖추고 있다.



Fig. 7. Project area

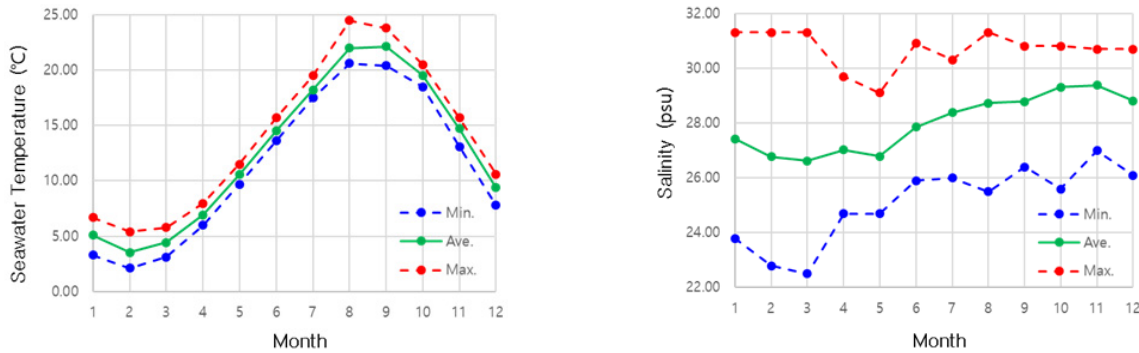


Fig. 8. Seawater temperature and salinity in Daesan (KHOA, 2003~2014)

3.2 기초자료 조사 및 EDA

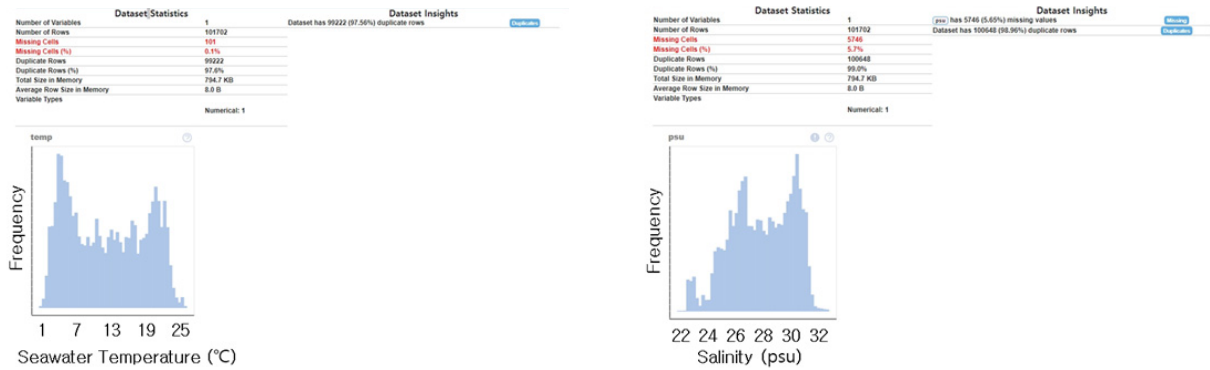
국립해양조사원(Korea Hydrographic and Oceanographic Agency, KHOA) 대산지역의 수온 및 염분도 월별 측정자료는 Fig. 8과 같다. 하절기 태풍 및 동절기 가뭄으로 해수 수온 및 염분도는 계절적인 특징을 가진다. 여름철은 수온이 높고 염분이 낮으며, 겨울철은 수온이 낮고 염분이 높게 나타나며, 수온은 2월에 2.5°C로 가장 낮고, 8월에 24.5°C로 가장 높았다. 염분은 3월에 22.50 psu로 가장 낮고, 1월에 31.30 psu로 가장 높았다. 국가 해양환경 정보통합시스템(MEIS, Marine Environment Information System) 해수 수온은 1.1~25.5°C, 평균 13.1°C로 나타났고, 염분은 평균 31.16 psu로 28.32~32.95 psu의 값을 나타냈다. 국립해양조사원(KHOA) 자료는 국가 해양환경정보통합시스템(MEIS) 자료와 비교하면 수온은 약 4%, 염분은 약 10% 낮았으며, 측정 장소가 항구 주변에 설치되어 지표수 유입에 대한 영향이 나타나는 것으로 검토되었다.

국가 해양환경 정보통합시스템(MEIS) 자료는 1997년부터 2020년(24년)까지 조사되어 가장 오랜 기간동안 측정된 값이 있지만, 연간 4회 자료가 측정되어 총자료의 수가 많지 않아

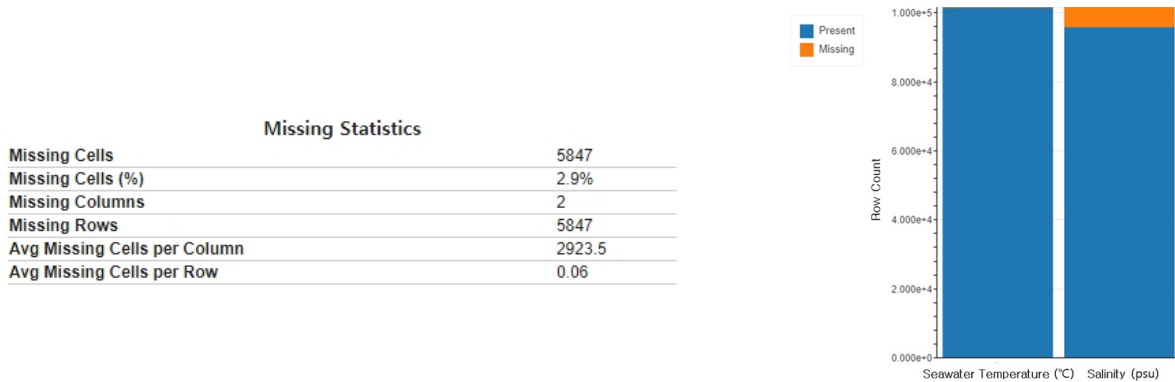
예측 모델의 입력자료로 사용하기에는 부족하다고 판단되었으며, 금회 모델은 자료는 2003년부터 2014년까지 1시간 간격으로 측정된 자료(약 10만건)가 있는 국립해양조사원(KHOA)의 실시간 데이터를 적용하였다. 주요 변수, 결측치 비중 등을 활용하기 위해 탐색적 데이터 분석(Exploratory Data Analysis, EDA)을 실시하였으며, 결측치를 제외한 자료의 개수는 수온 자료 101,601건, 염분도 95,956건으로 나타났다. 각 자료의 분포는 Fig. 9와 같고, 수온의 결측치는 101건(0.1%), 염분도의 결측치는 5,746건(5.7%)으로 나타나, 염분도 결측치가 수온에 비해 다소 많지만 전체 자료의 5% 수준이므로 모델 구축을 하는데 큰 문제는 없는 것으로 판단하였다.

결측치 분포를 상세하게 검토하기 위해 수온 및 염분도에 추가적인 분석을 실시하였으며, 결측치 비중은 Fig. 10과 같이 나타난다.

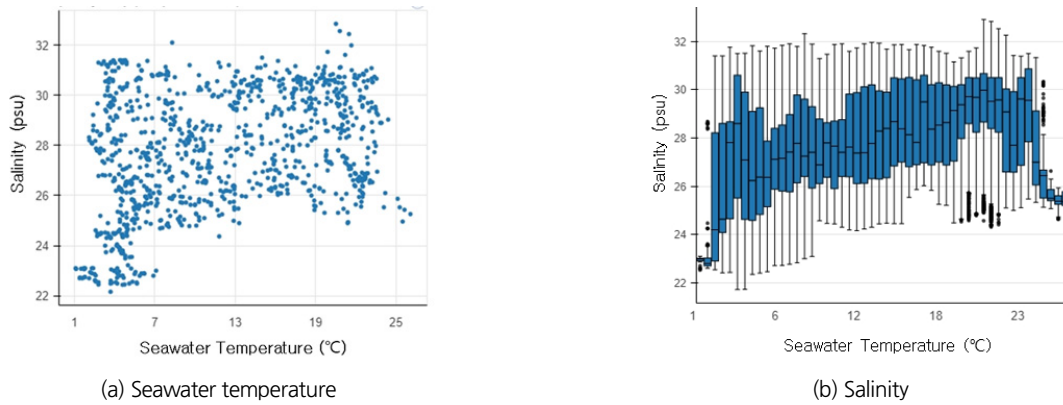
수온 및 염분도와 상관관계를 검토하기 위해 자료를 분석하였고, Fig. 11과 같이 수온 및 염분도는 큰 상관관계가 없는 것으로 검토되었다.



(a) Seawater temperature (b) Salinity
Fig. 9. EDA of seawater temperature and salinity (data distribution and missing data)



(a) Missing data analysis (b) Missing data distribution
Fig. 10. EDA of seawater temperature and salinity (missing data analysis)



(a) Seawater temperature (b) Salinity
Fig. 11. Scatter plot (left) and box plot (right) for seawater temperature and salinity

3.3 예측 모델 구축

대산지역의 2003년부터 2014년까지의 월별 수온 및 염분도는 Tables 1 and 2와 같다.

해양수질(수온, 염분도) 예측을 위해 딥러닝 모델을 구축하였으며, 구글에서 무료로 제공되는 Tensorflow를 활용하였

다. Tensorflow는 오픈소스 프로그램으로서 데이터 흐름 그래프를 사용하며, 그래프 노드는 수학 연산, 그래프의 선은 노드 간에 전달하는 데이터 배열(Tensor)를 나타낸다. Keras API를 통해 CPU 또는 GPU로 연산이 가능하며, Tensorflow 라이브러리를 기반으로 RNN-LSTM 딥러닝 모델을 구축하였다.

Table 1. Monthly seawater temperature in Daesan (2003~2014)

Year		Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
2003	Min.	1.6	2.5	3.4	5.7	8.5	12.5	15.7	19.2	20.2	15.8	10.6	7.0
	Ave.	4.5	3.4	4.4	7.17	10.8	14.3	17.6	20.6	21.1	18.5	14.5	9.3
	Max.	6.5	4.2	6.0	9.3	13.3	16.2	19.8	21.8	21.9	20.4	16.4	12.7
2004	Min.	3.3	3.1	3.2	5.2	8.8	12.7	15.5	20.1	19.2	17.1	13.0	7.3
	Ave.	5.51	3.9	4.6	7.6	10.8	14.6	17.6	20.9	20.4	18.7	15.1	10.6
	Max.	7.5	4.8	6.4	9.4	13.1	16.4	20.7	21.9	22.2	20.3	17.5	13.1
2005	Min.	1.0	1.8	2.0	4.1	8.1	12.1	15.7	19.7	20.3	17.2	12.2	4.4
	Ave.	5.3	2.6	3.12	6.1	10.1	14.2	18	21.5	21.5	19.45	15.1	7.8
	Max.	7.8	3.6	4.8	8.9	13.1	16.7	20.8	22.8	22.9	21.1	17.8	12.9
2006	Min.	3.0	1.7	1.5	4.6	7.8	11.7	15.5	19.1	20.6	18.7	12.8	7.3
	Ave.	4.3	2.9	3.8	6.4	10.1	13.9	17.5	21.9	21.6	20.2	15.7	10.5
	Max.	6.0	4.1	5.3	8.5	12.6	16.3	19.7	23.6	23.2	21.3	18.9	13.3
2007	Min.	5.3	4.7	4.7	6.6	9.4	13.5	16.9	20.4	19.2	19.4	7.7	7.0
	Ave.	6.7	5.4	5.8	7.97	11.5	15.7	18.9	22.4	22.6	20.5	13.1	9.9
	Max.	8.5	6.0	6.9	9.9	14.2	17.8	21.3	24.7	23.6	21.7	13.7	12.3
2008	Min.	3.7	2.3	2.9	5.8	9.0	12.7	15.8	19.8	20.7	18.0	12.4	7.3
	Ave.	5.6	3.2	4.5	7.5	11.1	14.6	18.09	22.0	22.9	20.4	15.0	10.2
	Max.	7.3	4.1	6.2	9.4	13.2	16.9	20.1	23.4	23.9	22.3	17.1	13.1
2009	Min.	3.9	3.5	3.9	5.7	8.8	12.7	16.9	21.8	22.4	18.8	12.3	5.8
	Ave.	5.46	4.4	5.1	7.6	11.1	15.1	19.5	22.7	23.1	20.4	15.2	9.7
	Max.	7.5	5.3	6.6	9.8	13.8	18	22.3	23.9	24.2	22.9	19.0	13.3
2010	Min.	3.5	2.6	3.6	4.8	7.1	11.2	15.2	20.5	22	16.6	10.9	4.6
	Ave.	4.7	3.4	4.4	6.4	9.9	13.9	17.8	22.5	23.8	19.7	14.0	8.4
	Max.	6.7	4.5	5.5	8.2	12.3	16.4	21.4	24.1	24.9	22.3	16.8	11.6
2011	Min.	0.8	1.2	2.3	4.5	7.5	11.3	15.1	20.2	20.4	16.8	12.9	6.7
	Ave.	3.3	2.1	3.3	6.0	9.7	13.6	17.8	21.9	22.2	18.6	15.6	9.8
	Max.	5.6	2.9	4.6	7.8	12	15.8	20.5	24.0	24.2	20.7	17.3	13.4
2012	Min.	3.3	1.7	2.5	4.4	7.7	12.1	15.7	20.3	20.1	16.5	11.1	6.1
	Ave.	5.2	2.6	3.5	6.17	10.4	14.6	18.0	21.2	21.1	18.7	13.9	8.8
	Max.	7.4	3.5	4.9	8.7	13.0	17.4	21.0	22.8	22.1	20.5	16.6	11.7
2013	Min.	4.3	3.5	4.4	5.0	7.4	11.2	15.7	21.3	21.2	17.0	10.1	5.8
	Ave.	5.5	4.6	5.5	6.41	9.9	14	18.6	24.5	22.8	19.1	13.85	8.7
	Max.	6.9	5.7	7.0	8.6	12.5	17.0	22.0	26.4	25.8	21.6	17.1	11.1
2014	Min.	4.0	3.2	4.0	6.4	9.2	13.4	17.3	20.1	21.7	17.9	13.9	6.5
	Ave.	5.2	4.0	5.2	7.9	11.5	15.7	19.3	21.7	22.4	20.0	15.5	9.0
	Max.	6.9	4.9	7.3	10.5	14.8	19.0	21.5	23.2	23.3	22.3	18.2	14.1

Python Language로 개발되었고, 각 패키지 모델은 딥러닝 모델을 구축할 때 적용되며, Numpy, pandas, seaborn은 데이터 가공 및 계산용, matplotlib은 수식을 그래프로 나타내기 위한 패키지이며, 모델 구축 시스템 사양 및 사용된 패키지 정보는 Tables 3 and 4와 같다.

구축된 모델은 시계열 자료 길이에 따른 가중치 손실을 막기 위해 RNN-LSTM 모델을 적용하였다. 전체 Data Set에서 Training Set (Training & Validation) 80%, Test Set 20%로 구분하여 학습을 진행하였고, 대산지역의 2003년부터 2014년까지의 수온 및 염분도 자료를 활용하여 장래 예측(6년) 수

Table 2. Monthly salinity in Daesan (2003~2014)

Year		Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
2003	Min.	31.0	31.0	30.1	28.4	27.6	27.4	28.1	28.7	27.9	28.8	29.1	29.2
	Ave.	31.3	31.3	31.3	29.7	28.5	29.3	30.1	29.8	29.1	30.1	30.4	30.5
	Max.	31.7	31.5	32.5	31.2	29.8	31.5	31.0	30.5	30.6	31.4	31.4	32.4
2004	Min.	29.5	28.8	27.4	25.8	25.1	26.0	25.9	28.1	29.1	28.0	28.0	28.6
	Ave.	30.6	30.2	28.8	26.9	26.27	26.9	27.7	30.0	30.3	30.0	29.9	30.0
	Max.	31.8	31.4	30.3	28.9	27.7	27.6	28.9	31.1	31.0	30.8	30.7	30.8
2005	Min.	28.0	27.8	26.2	24.6	24.6	25.6	27.55	29.5	28.5	29.2	29.7	29.6
	Ave.	29.3	29.0	28.3	25.5	25.1	25.9	27.75	29.6	29.9	30.8	30.7	30.7
	Max.	30.6	30.0	30.6	26.5	25.9	26.4	28.1	29.8	32.2	31.8	31.5	32.0
2006	Min.	26.8	27.3	25.9	25.5	25.2	24.6	24.0	23.4	25.8	27.45	29.1	27.9
	Ave.	28.5	28.3	27.3	26.2	25.6	25.9	26.0	26.3	26.82	28.56	30.3	28.9
	Max.	30.4	29.1	28.8	27.0	26.0	27.3	27.6	27.4	27.3	29.05	30.8	29.8
2007	Min.	26.7	25.8	25.5	25.0	26.2	26.1	27.6	27.2	26.8	29.0	27.55	29.45
	Ave.	27.9	26.7	26.5	27.5	27.2	28.29	29.4	29.3	30.0	30.5	28.36	30.84
	Max.	29.5	27.4	27.1	32.4	29.0	32.0	30.7	30.6	31.4	31.5	28.35	32.25
2008	Min.	23.3	23.8	22.7	22.6	26.6	29.1	28.4	27.8	28.2	28.5	26.8	24.4
	Ave.	24.9	24.5	23.7	24.7	29.1	30.9	30.3	30.5	30.6	30.3	29.4	26.54
	Max.	25.8	25.9	24.7	31.6	31.1	31.7	31.3	31.5	31.5	31.0	31.8	27.9
2009	Min.	23.7	22.8	22.3	27.1	25.8	25.8	22.9	28.4	29.6	28.4	26.6	26.6
	Ave.	24.9	23.3	25.6	28.6	26.8	26.4	27.7	29.9	30.8	30.7	30.5	30.7
	Max.	26.5	23.9	31.9	31.4	27.7	26.8	31.4	30.9	31.6	31.6	31.9	31.9
2010	Min.	29.4	25.9	25.2	27.0	27.0	27.2	26.8	26.7	25.0	24.3	25.5	24.7
	Ave.	30.6	27.5	27.6	28.2	27.9	27.9	28.0	28.6	26.9	25.6	27.9	27.0
	Max.	31.4	30.4	30.6	31.0	30.3	28.3	28.7	31.0	30.1	28.5	30.7	29.4
2011	Min.	22.8	22.5	20.96	22.37	25.74	26.02	25.3	25.44	25.83	25.89	25.39	25.3
	Ave.	23.8	22.8	22.5	27.4	26.6	29.0	26.8	26.7	26.7	27.0	27.0	26.09
	Max.	25.0	23.1	22.87	31.27	27.01	31.34	30.5	28.63	27.96	28.23	28.11	26.66
2012	Min.	24.49	24.09	24.27	23.28	23.56	29.09	28.5	28.38	27.0	26.28	25.86	26.38
	Ave.	25.2	24.4	27.0	25.0	27.8	30.4	29.9	31.3	29.6	29.0	28.5	28.4
	Max.	26.0	24.81	31.76	26.72	31.65	31.63	31.43	32.98	32.37	31.82	31.35	30.95
2013	Min.	24.45	23.92	24.2	24.45	24.12	24.21	24.48	24.56	25.71	26.21	26.82	27.77
	Ave.	25.6	24.7	24.8	26.8	24.7	26.4	26.8	25.5	28.3	28.6	29.0	29.0
	Max.	27.29	25.39	25.39	31.04	25.13	30.92	29.35	25.95	30.48	30.87	31.33	30.78
2014	Min.	25.57	25.57	25.29	25.41	25.32	25.38	28.01	26.44	25.5	26.53	26.75	26.97
	Ave.	26.5	28.6	26.0	27.7	25.8	27.0	29.5	27.2	26.4	29.85	29.6	29.1
	Max.	27.96	31.32	27.08	31.39	26.51	31.27	31.34	28.63	26.89	31.73	32.02	32.02

행하였다. 학습 횟수는 1000회로 설정하였으며, 오차함수로 mean_squared_error, 안정적인 학습을 위해 최적화(optimizer) 함수는 선행연구에서 많이 적용되며 일반적으로 사용되는 adam을 선정하였다(Xiao *et al.*, 2019). Activation 함수는 relu 를 사용하였으며, 모델 실행 이전에 성과를 저장하고, 5회 이

상 모델의 성과 향상이 없으면 최적화 단계에서 자동으로 학습을 중단하여 과적합을 방지하는 알고리즘(early_stop)을 적용하였다.

모델 성능 최적화를 위해 하이퍼파라미터 튜닝은 Test Size (1회 학습시킬 데이터 개수) 및 Win Size (예측 값 이전에 연속

Table 3. System specifications for model building

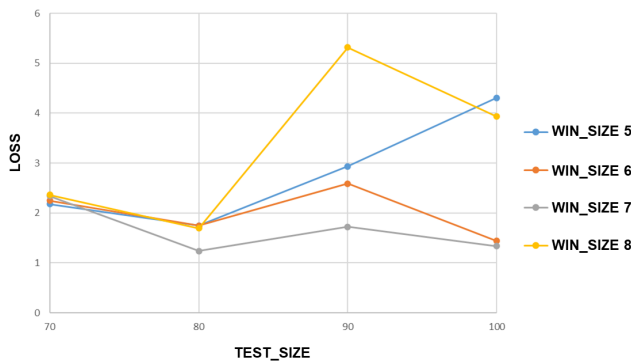
Item	Description
OS	Window 10
Programming Tool	Google Colab Pro
Language	Python 3.7.10
Deep Learning Library	Tensorflow 2.5.0
Deep Learning API	Keras 2.5.0
Algorithm	RNN-LSTM

Table 4. Python package information

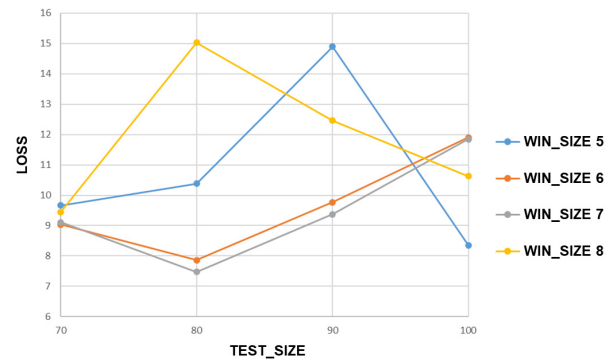
Package program	Version
numpy	1.19.5
pandas	1.1.5
matplotlib	3.1.1
seaborn	0.11.1

적으로 입력되는 자료 개수)를 검토하였으며, 수온에 대한 Loss 값은 1.24097~5.31512, 염분도는 7.47934~15.02981로 분포되었고, Fig. 12와 같이 TEST_SIZE는 80, WIN_SIZE는 7일 때 Loss 값이 가장 낮게 나타났다.

RNN-LSTM은 신경망 모델중에서도 시계열 예측에 뛰어난 것으로 알려져 있으며, 이전의 시간차를 이용해 미래를 예측하기 때문에 자료가 주기성(Seasonality) 및 자기상관성(Auto-correlation)이 높을수록 예측이 정확해진다. 예측된 값의 성능 비교를 위해 수정결정계수(adj. R²), 평균제곱근오차(Root Mean Square Error, RMSE), 절대오차(Mean Absolute Error, MAE)를 검토하였고, 금회 검토된 대산지역의 수온은 adj. R² 0.9398, RMSE 1.5989, MAE 1.1827, 염분도는 adj. R² 0.1151, RMSE 2.3429, MAE 1.9996로 계산되었다. Fig. 13(a)와 같이 수온은 일정한 주기성을 가지고 있지만, Fig. 13(b)와 같이 염분도는 태풍 및 집중호우 등에 의한 하천 방류량에 대한 영향

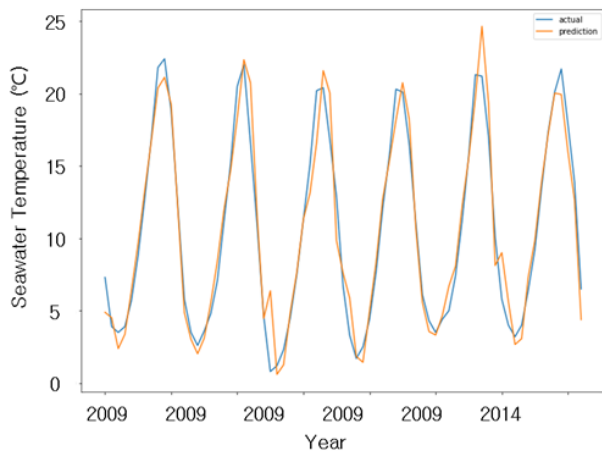


(a) Seawater temperature

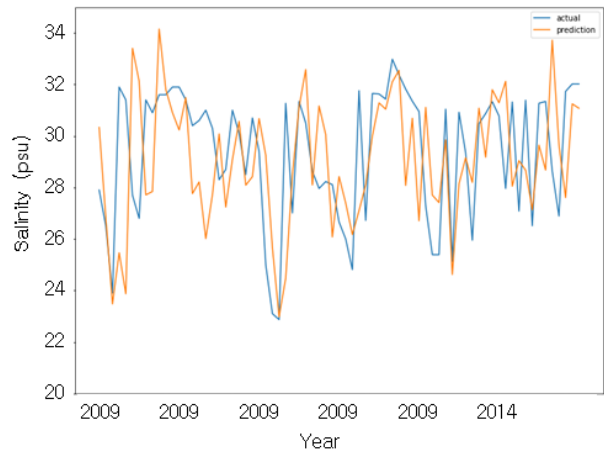


(b) Salinity

Fig. 12. Loss function graph



(a) Seawater temperature



(b) Salinity

Fig. 13. Prediction value graph

이 있어 예측 정확도가 상당히 낮아지고, 예측 값이 불규칙하게 나타났다. 그러나 향후 해안에서 먼 곳에서 해수가 취수된다면 하천 및 담수 영향이 적어지기 때문에 염분도 예측 정확도가 향상될 것으로 판단된다.

3.4 해수담수화 시설 전력비 비교

염분 및 수온 장래 데이터 예측값(2015~2020년)과 기존자료(2003~2014년)를 반영한 전력비 비교를 위해 물리화학식을 적용하여 검토하였다(Kim and Youm, 2019; K-water, 2019; Voutchkov, 2010). 주요 고려사항은 염분도, 수온, 유입수 성상, RO 공정구성 및 막 종류 등이 영향을 미치며, 염분도 및 수온 이외에는 설계단계에서 선정되기 때문에 염분도 및 수온에 대한 변화만 고려하였다. 세부적으로 적용된 내용은 Table 5와 같이 기존에 검토된 K-water (2019) 보고서를 참조하였으며, 대산해수담수화 시설의 용량은 일평균 기준 100,000 m³/d, 총 11개 계열(1개 예비)로 구성되어 한 계열 당 10,000 m³/d 으로 구성된다.

역삼투 공정에서 삼투압은 회수율에 따라 염분이 농축되면 증가한다. 농축된 염분도(TDSc, psu), 원수 염분도(TDSf, psu), 생산수 염분도(TDSp, psu) 및 회수율(Pr, %)로 나타낼 수 있고, Eq. (3)과 같다. 기존 대산지역 자료(2003~2014년) 원수 최대 염분도 32.98 psu, 회수율 50%, 생산수 염분도 0.2 psu를 적용하면, 농축된 염분도는 65.76 psu로 나타난다.

$$TDSc = \frac{TDS_f - TDSp \frac{P_r}{100}}{1 - \frac{P_r}{100}} \quad (3)$$

온도가 증가함에 따라 삼투압도 증가하게 되며, 해수의 삼투압(Po, bar), 기체상수(R, 0.0809(L·bar)/(mol·K)), 수온(T, °C), 해수 이온농도의 합(mi)으로 나타낼 수 있으며, Eq. (4)와 같다. 기존 대산지역 자료(2003~2014년) 최소 수온 0.8°C, 해

수 이온농도 2.0927(= 65.76 psu / 35 psu × 1.1135)를 적용하면, 삼투압은 46.35bar로 계산된다.

$$P_0 = R \times (T + 273) \times \sum m_i \quad (4)$$

역삼투압 공정에 소모되는 전력(P_{pump})은 Eq. (5)과 같이, 펌프 유량(Q, m³/min), 양정(H, m), 효율(η), 여유율(α, 0.1~0.15)로 계산될 수 있으며, 기존 대산지역 자료(2003~2014년) 1개 계열의 유량 6.9444 m³/min, 양정 463.5 m, 펌프효율 0.84, 여유율 0.15을 고려할 경우, SWRO 전력비는 1.72 kWh/m³로 계산되었다. 총 전력비는 SWRO 전력비의 약 45%(BWRO 전력비 30%, 기타 15%)를 할증한 2.50 kWh/m³로 검토되었으며(K-water, 2019), 장래 예측된 수질자료를 반영한 전력비는 Table 6과 같다.

$$P_{pump} = \frac{0.163 \cdot \gamma \cdot Q \cdot H}{\eta} \times (1 + \alpha) \quad (5)$$

기존 대산지역 2003년부터 2014년까지의 최소 수온은 0.8°C, 최대 염분도 32.98 psu이며, 금회 예측된 장래 최소 수온은 1.1°C, 장래 최대 염분도는 32.67 psu로 나타나 과거의 최소 또는 최대 값을 크게 벗어나지 않는 것으로 검토되었다. 해수담수화 전력비는 Table 6과 같이 기존 2.50 kWh/m³에서 금회 2.48 kWh/m³로, 약 0.02 kWh/m³가 감소 되었으며, 해수담수화에 필요한 전력은 장래에는 약 0.80% 줄어드는 것으로 나타났다.

현재 계획된 시설물은 적절한 것으로 판단되며, 장래 해수담수화 시설을 운영할 경우에는, 해수 수질변화를 고려한 시설물 운영방안이 필요할 것으로 검토되었다. 운영 최적화는 펌프의 효율을 최대화 할 수 있는 방안이 필요하기 때문에, 향후에는 해수담수시설 펌프 효율을 최적화 할 수 있도록, 더 많

Table 5. RO process (K-water, 2019)

Item	Unit	SWRO	BWRO	Remark	
Train	Type	EA	10+1	10+1	
	Flux	LMH	13.5	29.0	
	Recovery rate	%	52.6	90.0	Total 50%
	Membrane nos.	EA	9,240	4,620	
	Vessel nos.	EA	1,320	660	
Stage	Type	Stage	1 Stage	2 Stage	
ERD	Efficiency	-	97%	-	

Table 6. Power consumption comparison of existing and prediction value

Item	Existing data	Prediction data	Remark
Min. temperature (°C)	0.8	1.1	
Max. salinity (psu)	32.98	32.67	
Recovery rate (%)	50	50	
Pressure (bar)	46.35	45.95	
Power consumption (kWh/m ³)	SWRO	1.72	1.71
	BWRO	0.52	0.51
	Others	0.26	0.26
	Total	2.50	2.48

은 자료와 다양한 예측 모델을 적용한 연구가 필요할 것으로 판단된다.

4. 결론

본 연구를 통하여 머신러닝 기반의 전력비 예측을 통한 에너지 절감 방안을 모색하고, 기존 설계에 대한 대안을 제시하여, 해수담수화 시설의 규모가 과대 또는 과소하게 계획되는 것을 사전에 검토할 수 있었다. 해수담수화 공정에서 가장 많은 에너지를 소모하는 역삼투압공정의 전력비를 예측하기 위한 모델을 구축하고, 입력 자료는 국립해양조사원(KHOA)의 2003년부터 2014년까지 수집된 대산지역의 1시간 간격의 자료(약 105,000건)를 적용하였다. 수온은 월평균 수온은 최저 0.77°C, 평균 15.63°C, 최대 25.97°C이며, 월평균 염분도는 최저 24.12 psu, 평균 31.04 psu, 최대 32.98 psu로 나타났다.

전체 자료는 학습(Training Data, 80%) 및 테스트(Test Data, 20%) 자료로 구분하여, 주요 인자인 해수 수온 및 염분도에 대해 예측하였다. 장래 수온 및 염분도는 시계열의 자료 길이에 따른 가중치 손실을 막기 위해 RNN-LSTM 알고리즘을 적용하여 예측할 수 있는 머신러닝 기반의 모델을 구축했으며, 예측된 값은 테스트 자료와 검증하였다. 하이퍼파라미터 튜닝은 Loss 함수를 최소화하는 값을 비교하고, 최적화하는 방향으로 Test Size 및 Win Size 튜닝을 수행하였으며, Loss 값은 수온 1.24097, 염분도 7.47934로 나타났다.

장래 해양수질을 예측한 결과, 일정한 주기성 및 자기상관성이 높은 수온의 경우 예측이 상대적으로 잘되는 것으로 나타났다. 염분도는 폭우 및 태풍 등의 영향을 받아 불규칙한 패턴이 발생하여 Loss 값이 다소 높게 나타났다. 해수담수화 전력비는 염분도, 수온, 유입수 성장, RO 공정구성 등 전체적으로 고려하여 계산되었고, 과거자료 기준 2.50 kWh/m³, 금회 예측된 수질 기준으로는 2.48 kWh/m³로 나타나 해수담수화 전력비는 0.80%가 줄어드는 것으로 예측되었다.

최근 경영환경이 급격하게 변하고 있으며, ESG (Environment, Social, Governance)가 중요한 고려사항이 되고 있다. 향후 사회기반시설을 계획할 때에는 친환경, 사회적 책임 경영, 지배구조 개선 등 사회책임투자 및 지속 가능한 투자와 같이, 다양한 기준이 고려되어야 하며, 특히, 에너지를 많이 사용하는 건설 및 환경 플랜트는 이러한 요구는 점차 더 강해질 것으로 판단된다. 따라서, 전력비 최소화를 위한 운영 최적화는 앞으로도 풀어야 할 과제로 더 많은 연구가 필요하다. 본 연구에서 구축된 모델로, 해수담수화 시설의 운영전력비를 낮추면,

탄소 발생이 그만큼 줄기 때문에 운영자 입장에서는 에너지 절감으로 인한 경제적 이익뿐만 아니라, 탄소세 및 기업 이미지 제고 등의 부가적인 수익도 창출할 수 있을 것으로 검토되었다.

지금까지 장래 해수담수 시설물 규모에 대한 적절한 평가 방법 및 대안 제시가 없었기 때문에, 해수 온도 및 염분도 과거 자료를 활용하여, 역삼투압의 전력비를 예측하는 모델 개발 및 장래 해수의 수질 변화에 따른 전력비를 제안하였다는데 의의가 있다. 다만 본 연구에서 적용된 입력자료의 측정 기간이 짧고, 대산지역이 대산으로 제한된 점은 한계점으로 나타났다. 이번 연구에서는 장래 수질예측 모델을 구축하고 검토하는데 집중하였으며, 향후 연구에서는 이를 보완하여 다양한 지역의 해양관측소 자료 활용하고 장래 해수담수시설의 수집되는 수질 자료를 추가로 적용한다면 보다 신뢰성 있는 예측 모델을 구축할 수 있을 것이다.

머신러닝을 기반으로 한 빅데이터 분석을 산업 혁신과 변화에 적극적으로 적용하는 전략이 검토되어야 하는 시점에서 본 연구를 기반으로 다양한 예측모델 알고리즘을 검토하고 개발해 나갈 수 있을 것이다. 이러한 과정을 통해 해수담수 전력비 예측 모델의 정확도는 더욱 높아질 것이며, 시설물의 설계를 최적화하여 에너지 비용 절감에 기여할 것으로 기대된다.

감사의 글

본 연구논문은 서울과학종합대학원대학교 AI 빅데이터 MBA 학위 논문을 수정·보완하여 작성되었습니다.

References

- Badruzzaman, M., Voutchkov, N., Weinrich, L., and Jacangelo, J.G. (2019). "Selection of pretreatment technologies for seawater reverse osmosis plants: A review." *Desalination*, Vol. 449, pp. 78-91.
- Bernat, X., Gibert, O., Guiu, R., Tobella, J., and Campos C. (2010). *The economics of desalination for various uses. Re-thinking Water and Food Security: Fourth Botín Foundation Water Workshop*, CRC Press, London, pp. 329-346.
- Cho, H., Kim, D., Olivera, F., and Guikema, S.D. (2011). "Enhanced speciation in particles swarm optimization for multi-modal problems." *European Journal of Operational Research*, Vol. 213, No. 1, pp. 15-23.
- Choi, H.M., Kim, M.K., and Yang, H. (2020). "High water temperature prediction using RNN-based LSTM." *Proceedings of the Korea Institute of Information and Communication on*

- Conference, KIISE, pp. 693-695.
- Duda, R.O., Hart, P.E., and Stork, D.G. (2001). *Pattern classification*, Wiley-Interscience, New York, NY, U.S., pp. 3-11.
- Fritzmann, C., Löwenberg, J., Wintgens, T., and Melin, T. (2007). "State-of-the-art of reverse osmosis desalination." *Desalination*, Vol. 216, pp. 1-76.
- Global Water Intelligence (GWI) (2015). *Desalination markets 2016*. Oxford, UK., p. 93.
- Greenlee, L.F., Lawler, D.F., Freeman, B.D., Marrot, B., and Moulin, P. (2009). "Reverse osmosis desalination: Water sources, technology, and today's challenges." *Water Research*, Vol. 43, No. 9, pp. 2317-2348.
- Jeon, S.B., Oh, H.Y., and Jeong, M.H. (2020). "Estimation of sea water quality level using machine learning." *Journal of Korean Society for Geospatial Information Science*, Vol. 28, No. 4, pp. 145-152.
- Jeong, S., and Park, S. (2020). "Prediction of water temperature change using LSTM-based deep learning technology." *Proceedings of the Korea Institute of Information and Communication on Women's ICT Conference*, KIICE, pp. 42-43.
- Jun, G.I., Kwon, D.H., and Ki, S.J. (2020). "Comparing the performance of machine learning algorithms in predicting river water quality and quantity." *KSWST Journal of Water Treatment*, Vol. 28, No. 1, pp. 49-57.
- Jung, S., Kim, Y.J., Park, S., and Im, J.H. (2020). "Prediction of sea surface temperature and detection of ocean heat wave in the South Sea of Korea using time-series deep-learning approaches." *Korean Journal of Remote Sensing*, Vol. 36, No. 5-3, pp. 1077-1093.
- Kim, G.J., and Youm, K.H. (2019). "Design of dead-end membrane module with increased permeate flux by natural convection instability flow." *Korean Membrane Journal*, Vol. 29, No. 3, pp. 147-154.
- Kim, I.S., and Oh, B.S. (2008). "Technologies of seawater desalination and wastewater reuse for solving water shortage." *Korean Society of Environmental Engineers*, Vol. 30, No. 12, pp. 1197-1202.
- Kim, Y.H., Im, J., Ha, H.K., Choi, J.K. and Ha, S. (2014). "Machine learning approaches to coastal water quality monitoring using GOCI satellite data." *GIScience & Remote Sensing*, Vol. 51, No. 2, pp. 158-174.
- Kurihara, M., and Takeuchi, H. (2013). "Mega-ton water system: Japanese national research and development project on seawater desalination and wastewater reclamation." *Desalination*, Vol. 308, pp. 131-137.
- K-water (2019). *The feasibility and basic plan for SWRO project in Daesan industrial area (in Korean)*, pp. 6-80, 6-155, 6-179, 6-189, 6-192.
- Kwon, B.S., Lee, S.H., Kang, S.T., and Lim, J.L. (2020). "Current research trends and the need for localization in ultrapure water production facilities in semiconductor industries." *Korean Society of Environmental Engineers*, Vol. 42, No. 10, pp. 493-512.
- Lee, G.O., and Park, J.Y. (2020). "A case study on utilization of industrial wastewater reuse." *The Korean Society of Industry Convergence*, Vol. 23, No. 1, pp. 17-24.
- Lee, S.M., and Kim, I.K. (2021). "A comparative study on the application of boosting algorithm for Chl-a Estimation in the downstream of Nakdong River." *Journal of Korean Society of Environmental Engineers*, Vol. 43, No. 1, pp. 66-78.
- Mitchell, T.M. (1997). *Machine learning*. McGraw-Hill Science/Engineering/Math, New York, NY, U.S., pp. 81-108.
- Oh, I.S. (2008). *Pattern recognition*. Kyobo, pp. 1-15, 95-132.
- Park, J.S., and Lee, H.H. (2020). "Prediction of high turbidity in rivers using LSTM algorithm." *Journal of Korean Society of Water and Wastewater*, Vol. 34, No. 1, pp. 35-43.
- Rosenblatt, F., Stieber, A., and Shatz, R.H. (1957). *The perceptron a perceiving and recognizing automaton*, Report No. 85-460-1, Cornell Aeronautical Laboratory, Inc., New York, NY, U.S., pp. 1-2.
- Suh, S.H., Kim, K.W., Kim, H.H., Yoon, I.S., and Cho, M.T. (2015). "Evaluation of energy saving with vector control inverter driving centrifugal pump system." *KSFJ Journal of Fluid Machinery*, Vol. 18, No. 2, pp. 67-72.
- Tariq, R. (2016). *Make your own neural network*, Createspace Independent Publishing Platform, CA, U.S., p. 43.
- Voutchkov, N. (2010). *Desalination engineering: Planning and design*, McGraw-Hill, New York, NY, U.S., pp. 17, 73, 115.
- Voutchkov, N. (2018). "Energy use for membrane seawater desalination - current status and trends." *Desalination*, Vol. 431, pp. 2-14.
- Wikipedia (2021). Neuron, accessed 08 August 2021, <https://ko.wikipedia.org/wiki/%EC%8B%A0%EA%B2%BD_%EC%84%B8%ED%8F%AC>.
- Woo, S.W., and Kim, Y.H. (2017). "Reduction of power consumption by variable speed operation of high pressure pump in seawater reverse osmosis desalination plant." *KSFJ Journal of Fluid Machinery*, Vol. 20, No. 5, pp. 33-39.
- Xiao, C., Chen, N., Hu, C., Wang, K., Xu, Z., Cai, Y., Xu, L., Chen, Z., and Gong, J. (2019). "A spatiotemporal deep learning model for sea surface temperature field prediction using time-series satellite data." *Environmental Modelling & Software*, Vol. 120, 104502.
- Yang, Y., Dong, J., Sun, X., Lima, E., Mu, Q., and Wang, X. (2017). "A CFCC-LSTM model for sea surface temperature prediction." *IEEE Geoscience and Remote Sensing Letters*, Vol. 15, No. 2, pp. 207-211.
- Zhang, Q., Wang, H., Dong, J., Zhong, G. and Sun, X. (2017). "Prediction of sea surface temperature using long short-term memory." *IEEE Geoscience and Remote Sensing Letters*, Vol. 14, No. 10, pp. 1745-1749.