

## 인공지능 기반 온실 환경인자의 시간영역 추정

이정규<sup>1</sup> · 오종우<sup>2</sup> · 조용진<sup>3</sup> · 이동훈<sup>4\*</sup>

<sup>1</sup>충북대학교 바이오시스템공학과 대학원생, <sup>2</sup>티젟 테크놀로지 코리아 대리, <sup>3</sup>전북대학교 생물산업기계공학과 교수

<sup>4</sup>충북대학교 바이오시스템공학과 교수

## A Research about Time Domain Estimation Method for Greenhouse Environmental Factors based on Artificial Intelligence

JungKyu Lee<sup>1</sup>, JongWoo Oh<sup>2</sup>, YongJin Cho<sup>3</sup>, and Donghoon Lee<sup>4\*</sup>

<sup>1</sup>Graduate Student, Department of Biosystems Engineering, Chungbuk National University, Cheongju 28644, Korea

<sup>2</sup>Assistant Manager, Teejet Technologies, Inchon 21635, Korea

<sup>3</sup>Professor, Department of Bio-Industrial Machinery Engineering, Jeonbuk National University Jeonju 54896, Korea

<sup>4</sup>Professor, Department of Biosystems Engineering, Chungbuk National University, Cheongju 28644, Korea

**Abstract.** To increase the utilization of the intelligent methodology of smart farm management, estimation modeling techniques are required to assess prior examination of crops and environment changes in realtime. A mandatory environmental factor such as CO<sub>2</sub> is challenging to establish a reliable estimation model in time domain accounted for indoor agricultural facilities where various correlated variables are highly coupled. Thus, this study was conducted to develop an artificial neural network for reducing time complexity by using environmental information distributed in adjacent areas from a time perspective as input and output variables as CO<sub>2</sub>. The environmental factors in the smart farm were continuously measured using measuring devices that integrated sensors through experiments. Modeling 1 predicted by the mean data of the experiment period and modeling 2 predicted by the day-to-day data were constructed to predict the correlation of CO<sub>2</sub>. Modeling 2 predicted by the previous day's data learning performed better than Modeling 1 predicted by the 60-day average value. Until 30 days, most of them showed a coefficient of determination between 0.70 and 0.88, and Model 2 was about 0.05 higher. However, after 30 days, the modeling coefficients of both models showed low values below 0.50. According to the modeling approach, comparing and analyzing the values of the determinants showed that data from adjacent time zones were relatively high performance at points requiring prediction rather than a fixed neural network model.

**Additional key words :** indoor CO<sub>2</sub>, LSTM, prediction model, smart-farm, time complexity

## 서 론

우리나라 전체 인구 중에서 농업인구 비중은 1990년에 15.5%를 차지하였으나 2014년에는 5.5%로 줄어 인력이 부족한 상황이며, 농가소득은 2000년 23,072천원에서 2014년 34,950천원으로 연평균 3% 증가에 그치는 등 생산성 향상에 어려움을 겪고 있다(Kim 등, 2016). 또한 지역별 농업 인구 중 평균 50% 이상이 60세 이상의 농업인으로 고령화에 따른 노동력 감소는 농업 성장의 저해요소로 부각되고 있다(Statistics Korea, 2017). ICT를 농업의 생산, 가공, 유통 및 소비 전반에 접목하여 원격으로 관리 할 수 있는 스마트 팜을 보편적으로

확산하면, 노동 에너지 등 투입 요소의 최적 사용을 통해 우리 농업의 경쟁력을 한층 높이고, 미래 성장산업으로 걸인이 가능할 것으로 예측된다(Ju, 2017; Kim, 2016).

국내 스마트 팜 분야에 대한 연구는 주로 ICT를 활용하여 농작물 재배에 필요한 온도, 습도, 토양 영양상태 등을 각종 환경 계측 센서로 취득하여 상태를 파악한 후 이에 맞는 제어 장치를 구동하여 최적 환경을 구현하는데 중점을 둔 연구로 편향되어 있으며(Lee 등, 2016), 결과적으로 현재 국내 스마트 팜 기술은 시설 농업 선도국인 네덜란드에 9.8년의 기술적 격차로 뒤쳐져 있다고 분석되었다(Kim 등, 2016). 해외 농업 선진 국가들의 경우 2000년대부터 농가에서 취득한 빅 데이터를 정밀하게 분석하여 대상 작물에 따른 최적 환경 조건 개발을 통해 작물의 생산성과 품질 향상에 성공을 거두었다고 보고하였다(Hong 등, 2017). 우리나라의 경우 농업인의 경작 노하

\*Corresponding author: leedh@chungbuk.ac.kr

Received December 26, 2019; Revised June 18, 2020;

Accepted June 23, 2020

우에 의존한 관행적 재배 방식이 주를 이루었기 때문에 빅 데이터를 활용한 재배시스템 및 기술이 부족한 실정이다(Kwon 등, 2017). 2015년 이후 국가적 차원에서 개발 사업을 통해 ICT를 기반으로 한 지능형 스마트 팜 개발을 목표로 다양한 연구가 진행되고 있다(Kim 등, 2016).

온실 환경 최적 제어를 위한 연구에서 외부기상을 고려한 지능형 그린온실 시스템을 개발하여 각 장치들에 대한 온실 내 모델을 제안하고(Ooteghem, 2010), 난방 에너지 소비를 예측하고자 회귀모델과의 비교를 통해 타당성을 검증하는 등(Trejo-perca 등, 2009) 수학적 모델링에 기초한 연구에서, 복잡한 온실 내부 환경 변수들의 관계를 학습하여 변이 예측에 사용되는 인공지능 기법을 활용한 연구가 본격적으로 등장하였다(Choi 등, 2019). 온실 온도 예측을 위해 인공 신경망을 적용한 연구에서 제안한 모델과 회귀모델, 뉴럴 네트워크 회귀모델과의 비교를 통한 모델의 성능 평가가 진행되었다(Patil 등, 2008). 연속적으로 생성되는 온실 데이터 기반 자동 제어 신경망 모델 검증에 위해 다른 모델들과의 비교를 통해 NNAR(Neural Network Automatic Regression)모델의 성능이 우수하나, 데이터 측정에서 발생하는 측정 및 로깅 에러에 대한 별도의 검증 없이 모델을 구축하여 모델의 신뢰성을 보장하지 못한다는 단점이 있다(Patil 등, 2008). 전통적 방식의 온실은 기상조건에 따른 내부 환경 변화에 탄력적으로 대응하기 어렵기 때문에 RNN(Recurrent Neural Network)와 MLF(Multi Layer Feedforward) 방법을 이용해서 반대 역학

모델을 학습하고 예측 모델의 시뮬레이션을 통해 타당성을 평가 비교하였으며, RNN이 복잡한 온실 환경 제어에 적합한 모델이라고 분석되었다(Fourati와 Chtourou, 2007). 한편 실시간으로 실내 환경을 예측하는 방법은 RNN 모델에 비해 LSTM(Long Short Term Memory)의 예측 성능이 140% 개선된 것으로 나타났다(Kim과 Oh, 2018).

본 연구에서는 실시간으로 실내 환경 변화를 예측할 수 있는 방법으로 최상의 결과를 나타낸 LSTM을 활용하여 온실 내부 CO<sub>2</sub>의 변화를 예측하기 위한 두가지 방법의 모델링을 수립하고 예측 성능을 평가하였다.

## 재료 및 방법

### 1. 계측 시스템

온실 내부 공간의 환경인자를 계측하기 위한 모듈의 센서부는 디지털 온·습도 센서(HDC1080, Texas Instruments, USA), 조도 센서(TSL2561, TAOS, China), CO<sub>2</sub> 센서(TG100, TRUEYES, Korea)를 사용하였다. 센서 모니터링 시스템을 구성하기 위해 각 환경인자 계측 모듈을 Fig. 1과 같이 제작하였다. 고온 다습한 온실 환경에서 내구성을 위하여 플라스틱 케이스를 사출하였고, 내부에 온·습도, 조도, CO<sub>2</sub> 센서를 설치하였다. CPU(CC1352, Texas Instruments, USA)는 RF(Radio frequency)무선 송수신 장치와 ARM(Acorn RISC machine)코어를 내장한 3.3V의 저전력 프로세서이다. CPU

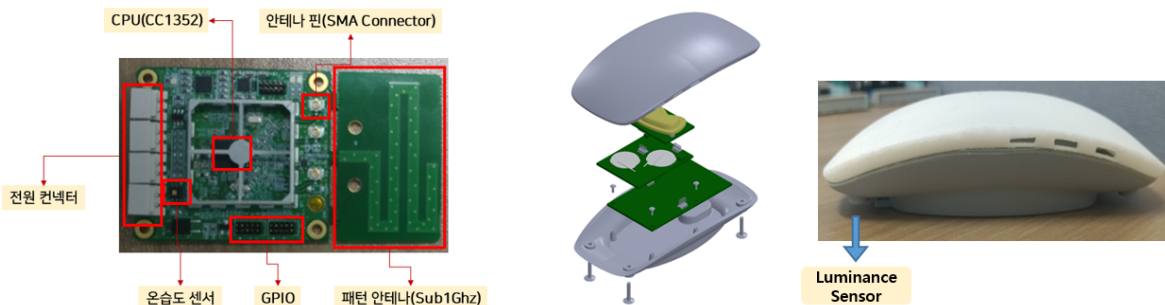


Fig. 1. Environmental factor measurement module and circuit diagram.

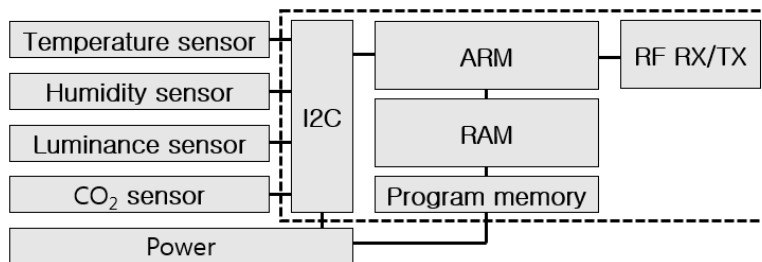


Fig. 2. Hardware configuration diagram of environmental factors measurement.

와 온 습도, 조도, CO<sub>2</sub> 센서의 값은 I2C 라인으로 데이터와 제어 신호를 통신한다(Fig. 2).

실험을 위한 데이터는 딸기재배 플라스틱 온실(300m<sup>2</sup>, 충청북도 청주시 상당구 가덕면 노동리 39)에서 확보하였다. 해당 농가에서는 설향 품종의 딸기(*Fragaria x ananassa* Duch.)를 재배하였으며 폭 10m, 길이 100m, 측고 1.2m, 동고 3.6m의 규격의 단동 아치형 플라스틱 온실이다. 실험을 실시하는 기간 동안 특별히 보온이 필요한 경우를 제외하고 측창을 오픈한 상태의 자연환기방식의 환경관리를 실시하였다. 온실 내 각 위치에 따라 환경 요인 측정값이 서로 다르기 때문에 앞서 제작한 환경인자 계측 모듈 총 18개를 25m 간격으로 3개 구역에 지상에서 1m, 2m 의 높이에 각 6개씩 설치하여 2018년 3월 1일부터 4월 30일까지의 60일 간의 온실 내부의 환경데이터를 수집하였다(Fig. 3). 계측된 센서 데이터의 값은 초당 1회씩 무선으로 농장에 설치된 노트북으로 전송할 수 있도록 하였다.

## 2. 데이터 전처리

데이터 분석 작업을 위해 정확한 센서 데이터 값이 전제되어야 한다. 따라서 본 연구에서는 데이터를 정제하기 위해 데이터 결측치는 보간법으로 보완하였으며, 노이즈가 심한 CO<sub>2</sub> 데이터는 0~1600μmol·mol<sup>-1</sup>을 기준으로 평활화(Smoothing) 과정을 통해 데이터를 전처리(Preprocessing)하였다. 결손된 데이터는 평균값을 적용하기 위해 moving average를 사용했다(Zhang 등, 2018). 또한 센서 데이터의 편차가 매우 넓기 때문에 이를 동일하게 맞춰주기 위해 3점 이동 평균인 (1)식을 이용하여 계산 후 보정하였다.

$$x_i = \frac{x_{i-1} + x_i + x_{i+1}}{3} \quad (1)$$

$x_i$  = Actual measurement value.

$x_{i-1}$  = Previous actual measurement value.

$x_{i+1}$  = Subsequent actual measurement value.

정규화(Normalization)는 데이터의 최대, 최소값을 이용하여 입력 변수의 값을 0부터 1사이의 값으로 변환시켜 주는 것으로 상대적인 크기를 비교하기 위하여 데이터의 범주를 바꿀 때 사용한다(Jo, 2019). 스마트 팜에서 수집된 데이터는 단위들이 각각 다르고, 그 편차가 매우 크기 때문에 (2)식을 이용한 정규화 과정으로 측정 데이터를 변환하였다.

$$x_{normal} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

$x_{normal}$  = Normalization value of measured data.

$x_{max}$  = Maximum value of measured data.

$x_{min}$  = Minimum value of measured data.

## 3. LSTM

RNN을 이용하면 시계열 데이터를 효과적으로 모델링 할 수 있지만 시간에 따라 학습이 계속 진행되면, 이전에 입력층에서 받은 정보가 학습에 미치는 영향이 점점 감소하다가 사라지게 된다. RNN의 은닉층을 입력 게이트, 출력 게이트, 기억 게이트라는 세 가지 게이트로 구성된 메모리 블록(Memory Block)으로 대체한 구조를 지닌 LSTM 네트워크가 제안되었다(Gers 등, 2000). Fig. 4에서 T에서의 입력 데이터를 받은 이후에 입력 게이트를 닫아버려서 새로운 입력값을 받지 않고, 기억 게이트를 열어놔서 T에서의 입력 데이터를 계속해서 전달받으면, T에서의 입력의 영향력을 계속해서 가져갈 수 있다. 마지막으로 출력 게이트를 열고 닫으면서, T에서의 입력 데이터의 영향력을 반영하거나 반영하지 않을 수 있다. 식 (3)은 LSTM의 수학적 공식이다.  $c_t$ 는 셀 상태 벡터이고,  $W, U, b$ 는 각각 파라미터 행렬 및 벡터이다.  $f_t, i_t, o_t$ 는 게이트 벡터들이다.  $f_t$ 는 Forget 게이트 벡터로서 이전 셀 상태를 기억하는 가중치 역할을 한다.  $i_t$ 는 입력 게이트 벡터로서 새로운 정보를 획득하는 가중치 역할을 한다. 반면에,  $o_t$ 는 출력 게이트 벡터이며, 출력 후보를 선택하는 역할을 한다.  $x_t$ 는 입력 벡터이고,  $h_t$ 는 출력 벡터이다. LSTM에서 2 종류의 활성화 함수가 사용된다.  $\sigma$ 는 Sigmoid 함수이고,  $\tanh$ 는 하이퍼볼릭 탄젠트 함수이다(Kim과 Oh, 2018).

$$\begin{aligned} f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \bar{C}_t &= \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \\ C_t &= f_t * C_{t-1} + i_t * \bar{C}_t \\ o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t &= o_t * \tanh(C_t) \end{aligned} \quad (3)$$

## 4. 모델링 설정

센싱한 환경인자 데이터는 전처리 과정 후 도출된 시계열 데이터를 학습화하기 위해서 실측값을 분할하는 시퀀스 리그리션 과정이 필요하다. 데이터 처리는 T시점 평균값과 T+1의 평균값으로 분류한 뒤 T+2의 시점을 실측 결과값을 예측하였다. 같은 과정을 시계열 데이터에 따라 반복하여서 60일치의 시

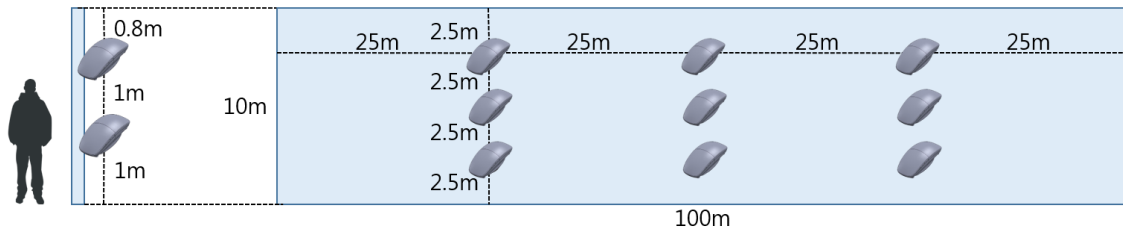


Fig. 3. The floor plan with location of the sensors for experiment.

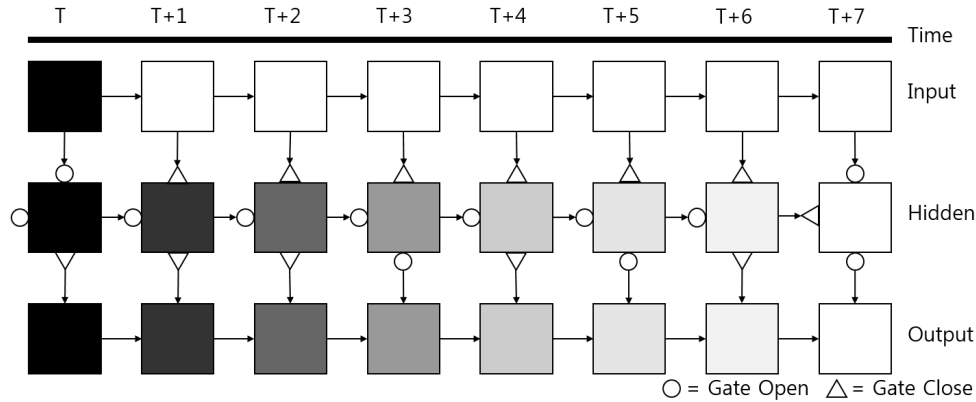


Fig. 4. Solve the problem of (t) losing its influence of the brightness of the shadow with LSTM (Long Short Term Memory).

간 단위 별로 분류하였다. 학습을 위한 데이터는 분 단위의 평균값을 사용하였다. 데이터를 예측하는 하는 방법을 예로 들면, T 시점 0시 1분 데이터와 T+1 시점의 0시 2분의 데이터가 학습의 과정을 거친 후 T+2 시점 0시 3분의 데이터를 예측하고 정확도를 결정계수로 표현한다. 분 단위 평균 데이터는 각 시점의 차이가 존재하므로 각 1438개의 데이터 셋으로 변환하였다. 시퀀스 리그리션 과정을 거친 데이터는 Matlab 2014 (Mathworks, USA)의 LSTM 네트워크를 사용하여 학습화하였고, LSTM 기본 수식 기초로 학습을 진행해서 결과값을 나타낸다. 시계열 데이터의 정확도를 높이기 위해 lstm\_layer를 복층으로 설계하고 MaxEpochs도 200회로 설정하여 모델링 1과 2를 단계별로 처리하였다. 모델링 1의 input은 60일의 평균데이터, 모델링 2의 input은 예측하고자 하는 날의 전일 데이터를 표본으로 학습을 진행하여 예측값의 정확도를 비교 분석하였다. Fig. 5(A)는 모델링 1로, 수집한 60일의 데이터의 평균으로 하루의 예측 모델을 설정하여 나머지 데이터 총 60개를 학습하는 방식을 도식화한 것으로 실험에서는 실측값과 비교한 결과를 도출하였다. Fig. 5(B)는 모델링 2로, 시계열에 따라 모델링이 진행되는 과정을 보여준다. 1일차 학습한 모델링을 통해 2일차 예측값을 만들고 실제 실측값과 비교한다. 2일차 학습에는 2일차만 데이터를 학습하여 모델링 후 나온 3일차 예측값을 실측값과 비교한다. 이러한 방식으로 현재의

예측값은 전날의 모델링을 통해서만 예측하도록 하고 실측값과 비교한 결과를 도출하였다.

## 결과 및 고찰

인공 신경망 모델에 학습하기에 앞서 필요한 입력변수의 현황을 고려하기 위하여 실험기간 스마트 팜 내 2019년 3월 1일부터 4월 30일까지 총 60일의 기본적인 환경 데이터 CO<sub>2</sub>를 분석하였다. Fig. 6은 총 60일 간의 초당 취득된 CO<sub>2</sub> 데이터를 일 평균값으로 나타내었다. 전체 평균 CO<sub>2</sub> 최대값은 337 μmol·mol<sup>-1</sup>, 최소값은 253 μmol·mol<sup>-1</sup> 이며, 평균 CO<sub>2</sub>값은 294 μmol·mol<sup>-1</sup> 으로 나타났다.

모델링 방식을 다르게 하여 예측값과 실제 CO<sub>2</sub> 값을 비교 분석하였다. Fig. 7은 모델링 학습 방식에 따른 CO<sub>2</sub> 결정 계수를 비교한 그래프이다. 두 가지 모델링이 형태적으로 유사하게 나타났다. 1일부터 30일 사이 예측 성능이 좋은 구간에서는 전날 학습을 통해 예측을 하는 모델링 2의 결정계수가 모델링 1에 비해 약 0.05 정도 높게 나타났다. 특정 일자를 제외하고 0.70 ~ 0.88의 결정계수를 나타내어 예측 성능의 신뢰도를 확인할 수 있었다. 하지만 30일 이후에는 두 가지 모델링 모두 결정 계수 값이 0.50 이하로 낮은 값을 보였다. CO<sub>2</sub>는 다른 나머지 환경인자들에 비해 환기, 작물의 생육 등으로 인해 학습

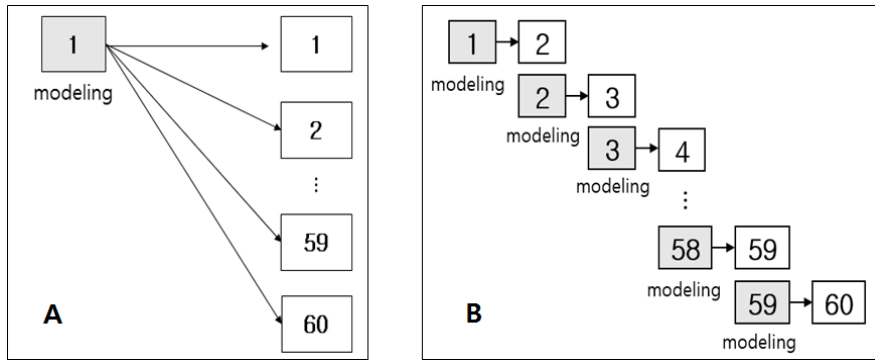


Fig. 5. The CO<sub>2</sub> prediction in indoor-greenhouse using artificial intelligence (A: Typical model for modeling; Modeling 1, B: Time series modeling; Modeling 2).

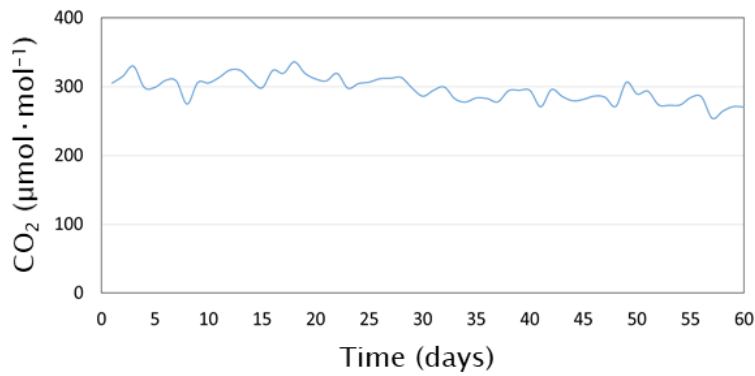


Fig. 6. Average CO<sub>2</sub> a day during 60 days at the experimental greenhouse using 18 sensors.

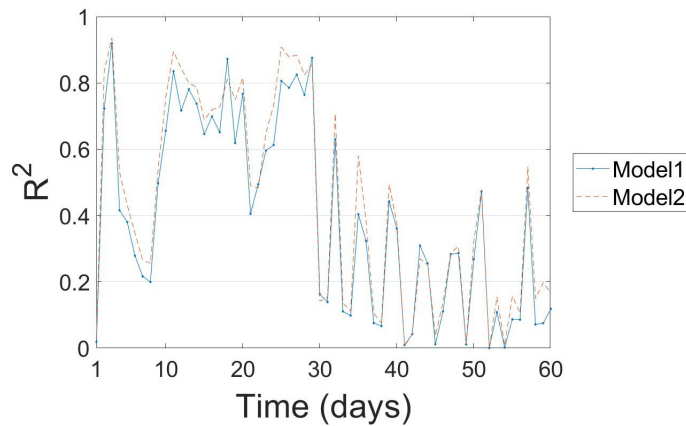


Fig. 7. Comparison plot of determination coefficient resulted from the model 1 and 2 to predict CO<sub>2</sub> (Model1: Typical model for modeling, Model2: Time series modeling).

에 영향을 끼치는 요소들이 많아 예측이 어려웠다. 30일 이후 점차 센서 모듈에서 측정값의 노이즈가 심해지는 것을 확인할 수 있었다. 플라스틱 온실은 외부 환경 변이에 의한 내부의 불균일성이 존재하며(Vranken 등, 2005) 재배기간 특별한 제어를 하지 않아 외부환경에 노출되었고, 30일 이후 일 평균 농도가  $300\mu\text{mol}\cdot\text{mol}^{-1}$  이하에서 불규칙한 변화(Fig. 6)를 나타냈

던 점, 머신러닝 기반 예측모델 비교실험에서도 나타났던 하루 및 계절별 예측모델의 불안정성(Kim 등, 2018)의 원인으로 작용하여 30일 이후 예측 성능이 감소한 것으로 판단된다(Fig. 7).

Fig. 8은 60일 중 첫 날을 제외한 가장 높은 결정계수 값(0.88)을 보인 28일의 실측 데이터와 예측 데이터를 비교분석

한 모델링 2의 그래프이다. 실제 센서에서 취득된 값과 예측된 값의 형태적으로 일치하였다. 전반적으로 예측 성능이 우수했던 30일 이전의 CO<sub>2</sub>의 일간 변화는 일정한 형태를 보여주고 있다.

Fig. 9는 60일 중 가장 낮은 평균 결정계수 값(0.31)을 보인 구간 47, 48, 49일 모델링 2의 CO<sub>2</sub> 그래프이다. 낮은 예측값이 나온 48일의 전날 47일 CO<sub>2</sub> 데이터에서 새로운 형태의 값들이 취득되었다. 4월에 접어들면서 내부 온도를 낮추기 위한 환기 또는 센서 노이즈 값이 심해진 것으로 유추된다. 30일 이후에는 평균적인 일정한 형태와는 다른 값들 학습에 적용되었고, 대부분이 낮은 결정계수 값으로 예측되었다. 광도와 온도 가 높을 경우 식물의 광호흡량이 증가하지만, 광합성도 증가

하여 결국 내부 CO<sub>2</sub> 농도는 낮아진다(Jung 등, 2017; Park 등, 2010). 30일 이후 시점의 날씨는 4월 1일 이후로, 외부 기온의 상승 등 외부 환경조건의 변화로 인하여 예측 모델에 영향을 미친 것으로 판단된다. 판단되며 예측 성능이 우수한 Fig. 8의 농도 변화와 다른 모형의 그래프로 나타났다.

평균 제곱근 오차를 모델링 방식에 따라 구별한 결과, 모델링 1의 경우 결정계수에 비례하여 오차값이 모델링 2에 비해 상대적으로 높고, 오차 범위 최대값은 70 μmol·mol<sup>-1</sup>, 최소값은 18 μmol·mol<sup>-1</sup>을 나타내었다. 모델링 2는 오차값의 최소 오차 범위는 10 μmol·mol<sup>-1</sup>을 나타내었다. CO<sub>2</sub> 결정계수 그래프에서 상관도가 낮았던 지점에서 동일하게 오차값 또한 높게 나타났다. 모델링 2가 상대적으로 모델링 1에 비해 낮은 오차

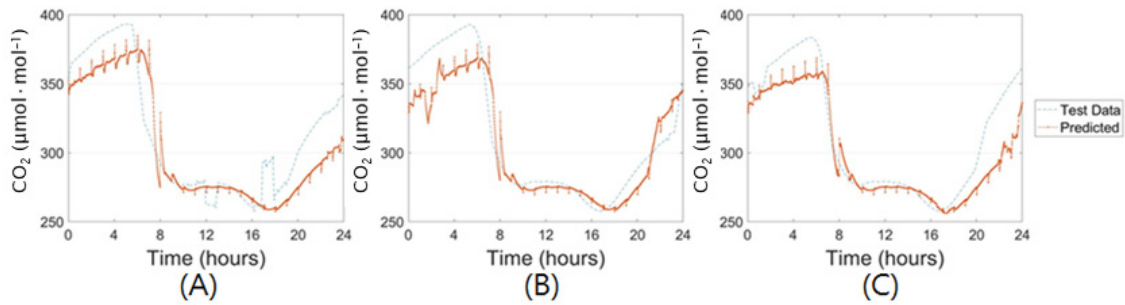


Fig. 8. Comparison plot of measured and estimated CO<sub>2</sub> at (A) 27<sup>th</sup>, (B) 28<sup>th</sup> and (C) 29<sup>th</sup> day when show the highest determination coefficient.

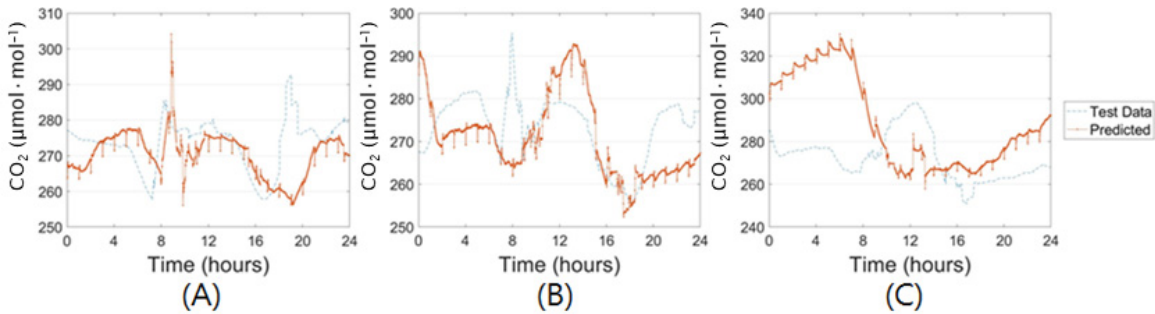


Fig. 9. Comparison plot of measured and estimated CO<sub>2</sub> at (A) 47<sup>th</sup>, (B) 48<sup>th</sup> and (C) 49<sup>th</sup> day when show the lowest determination coefficient.

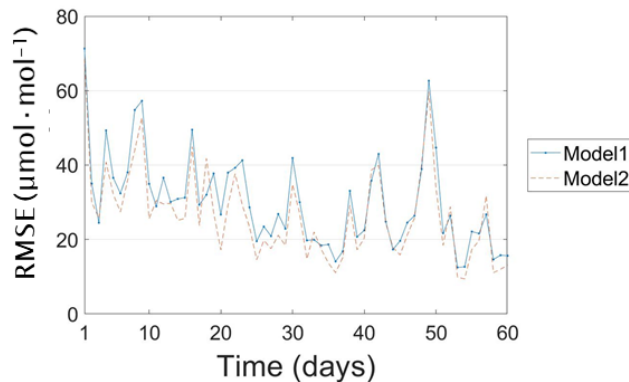


Fig. 10. Comparison of CO<sub>2</sub> prediction error rate by modeling (Model1: Typical model for modeling, Model2: Time series modeling).



을 보였다. 두 모델링 모두 60일 동안 대체적으로 오차값은  $30\mu\text{mol}\cdot\text{mol}^{-1}$ 을 유지하였다(Fig. 10).

결과적으로 모델링 접근법에 따라 분석 한 결과, 오늘의 데이터로 내일을 예측하는 모델링 2가 전체 평균치를 기준으로 내일을 예측하는 모델링 1보다 예측 성능이 우수하였다. 예측 성능이 우수한 시점의 일간  $\text{CO}_2$ 의 변화추세는 야간부터 6시까지 상승하며 일출 이후에 급격히 하락하는 것을 확인할 수 있다. 이는 야간에는 호흡으로 인한 배출, 주간에는 광합성 호흡에 의한 흡수 작용에 의해 내부 농도의 변화에 일정한 흐름이 있음을 예상할 수 있다. 다만 기상이 급변하거나 온실제어가 일상적이지 않은 날 또는 계절의 변화로 외부 환경이 작물의 성장 여부에 따른 온실 내의  $\text{CO}_2$  분포가 달라지는 경우, 예측 성능은 감소하며 일간  $\text{CO}_2$  농도의 흐름도 다른 추세를 나타내게 된다. 본 연구에서는 외부 환경 변수의 영향을 고려하지 않은 상태에서 진행하였기 때문에 후속 모델링 생성에는 환경 변수에 대한 상관관계의 분석이 필요하며, 최적화된 모델링 방법의 개선이 필요하다. 또한 작물에 대한 성장 인자 데이터를 추가하면 스마트 팜에서 불확실한 내부 환경 변화의 예측 가능성이 높아질 수 있을 것으로 예상된다.

## 적 요

스마트 팜 관리의 활용 효율성을 높이기 위해서는 작물 및 환경 변화에 대한 사전 검사를 실시간으로 평가하기 위한 모델링 기법이 필요하다. 시설 온실 내부의  $\text{CO}_2$ 와 같은 필수 환경 요소는 다양한 상관 변수가 밀접하게 결합된 시간 영역에서 신뢰할 수 있는 추정 모델을 확립하기가 어렵다. 따라서 본 연구는 입력 영역과 출력 변수를  $\text{CO}_2$ 와 같은 시간 관점에서 인접 영역에 분포된 환경 정보를 이용하여 시간 복잡도를 줄이기 위한 인공 신경망을 개발하기 위해 수행되었다. 스마트 팜을 계측하기 위한 센서 모듈을 통해 환경 요소를 지속적으로 측정하였다. 실험기간의 평균 데이터로 예측하는 모델링 1, 전일 데이터로 예측하는 모델링 2를 구성하여  $\text{CO}_2$  환경인자의 상호관계를 예측하였다. 전일의 데이터 학습으로 예측하는 모델링 2가 60일 평균값으로 예측한 모델링 1에 비해 성능이 우수하였다. 30일 이전까지는 대부분 0.70~0.88사이의 결정계수를 보였으며 모델링 2가 약 0.05정도 높게 나타났다. 하지만 30일 이후에는 두 가지 모델링 모두 결정 계수 값이 0.50 이하로 낮은 값을 보였다. 모델링 접근법에 따라 결정요인의 값을 비교하고 분석한 결과 인접한 시간대의 데이터는 고정 신경망 모델을 사용하는 대신 예측이 필요한 지점에서 상대적으로 높은 성능을 나타냈다.

**추가 주제어:** 기계학습, 스마트 팜, 시간 복잡성, 환경예측

## 사 사

본 연구는 농촌진흥청 연구개발과제(PJ0139252019)와 국립식품기술기획평가원의 첨단생산 기술개발사업(319013-01-1-HD020)의 지원으로 수행되었음.

## Literature Cited

- Choi, H.Y., T.W. Moon, D.H. Jung, and J.E. Son. 2019. Prediction of air temperature and relative humidity in greenhouse via a multilayer perceptron using environmental factors. Protected horticulture and plant factory. 28:95-103 (in Korean).
- Fourati, F. and M. Chtourou. 2007. A greenhouse control with feed-forward and recurrent neural networks. Simulation modelling practice and theory. 15:1016-1028.
- Gers, Felix.A., Jrgen. Schmidhuber, and Fred. Cummins. 2000. Neural Computation. 12:2451-2471.
- Jo, J.M. 2019. Effectiveness of normalization pre-processing of big data to the machine learning performance. Journal of the KIECS. 14:547-552 (in Korean).
- Ju, H.J. 2017. The model evaluation of the greenhouse temperature management and the temperature management scenario to warm root zone during low temperature season for smart farm. Journal of Agriculture & Life Sciences. 48:13-19 (in Korean).
- Jung, D.H., H.I. Yoon, and J.E. Son. 2017. Development of a three-variable canopy photosynthetic rate model of romaine lettuce (*Lactuca sativa* L.) grown in plant factory modules using light intensity, temperature, and growth stage. Protected Horticulture and Plant Factory, 26:268-275 (in Korean).
- Kim, J.H. 2016. 4th industrial revolution, education in the era of artificial intelligence. STSS sustainability science conference. 21-29 (in Korean).
- Kim, S.K., Y.G. Cha, and G.S. Kim. 2016. Smart farm will lead future farming. Samsung KPMG KRI research institute. Issue monitor (in Korean).
- Kim, S.K., and T.I. Oh. 2018. Real-time PM10 concentration prediction LSTM model based on IoT streaming sensor data. Korea academy industrial cooperation society. 19:310-318 (in Korean).
- Kim, S.Y., K.S. Park, S.M. Lee, N.M. Heo, and K.H. Ryu. 2018. Development of prediction model for greenhouse control based on machine learning. Journal of digital contents society 19:749-756 (in Korean).
- Kwon, H.M., S.E. Hong, D.S. Park, H.J. Kim, and Y.G. Kim. 2017. A study on the analysis of major crop models for optimum growth conditions. Proceedings of the korean institute of communication sciences conference. 1539-1540

- (in Korean).
- Lee, J.S., Y.G. Hong, K.H. Kim, D.H. Im, and S.R. Han. 2016. A study on the development of cloud system for automatic control of smart greenhouse. Proceedings of the Korean institute of communication sciences conference. 559-560 (in Korean).
- Ooteghem, R.J.C. 2010. Optimal control design for a solar greenhouse. IFAC proceedings 43:304-309.
- Park, S.A., M.G. Kim, M.H. Yoo, M.M. Oh, and K.C. Son. 2010. Comparison of indoor CO<sub>2</sub> removal capability of five foliage plants by photosynthesis. Korean journal of horticultural science & technology 28:864-870 (in Korean).
- Patil, S.L., H.J. Tantau, and V.M. Salokhe. 2008. Modelling of tropical greenhouse temperature by auto regressive and neural network models. Biosystems engineering 99:423-431.
- Statistics Korea. 2017. KOSIS national statistical portal. Available at: <http://www.kosis.kr>. accessed september 21, 2018.
- Trejo-Perea, M., G. Herrera-Ruiz, J. Rios-Moreno, R.C. Miranda and E. Rivas-Araiza. 2009. Greenhouse energy consumption prediction using neural networks models. International journal of agriculture & biology. 235-249.
- Vranken, E., R. Gevers, J.M. Aerts and D. Berckmans. 2005. Performance of model-based predictive control of the ventilation rate with axial fans. Biosystems Engineering. 91:87-98.
- Zhang, Xiao., Ling. Xu, Feng. Ding, and Tasawar. Hayat. 2018. Combined state and parameter estimation for a bilinear state space system with moving average noise. Journal of the franklin institute. 355:3079-3103.