

다중 회귀 분석을 이용한 파프리카 생산량 예측

황인철*, 노희선*, 양동일**, 김만배^o

Prediction of Paprika Yield Using Multiple Linear Regression

In-Chul Hwang*, Heesun Noh*, Dong-Il Yang**, Manbae Kim^o

요약

최근 스마트팜을 구축하여 빅데이터와 IoT 기술을 활용한 농작물의 생산성 향상에 주력하고 있다. 이에 필요한 생산량 예측은 디지털 농업에서 중요한 분야로 다양한 방법들이 활용되고 있다. 본 논문에서는 스마트팜 파프리카 농가에서 수집된 환경 및 생산량 데이터를 이용한다. 먼저 다중회귀분석을 이용하여 환경데이터와 생산량의 최적 모델을 구축하고, 이것을 기반으로 생산량을 예측하는 방법을 제안한다. 실험에서는 예측 생산량의 적합도를 조사한 결과, R^2 값은 [0.9171, 0.9714]의 분포를 가져, 만족스러운 예측 성능을 얻었다. 제안 방법은 의사결정시스템 구축에도 활용이 가능하다.

키워드 : 생산량 예측, 최적 모델, 스마트팜, 회귀분석, 상관계수

Key Words : yield prediction, smart farm, optimal model, regression analysis, cross correlation

ABSTRACT

Recently, in agriculture, high crop productivity is a main aim that can be achieved by integrating bigdata and IoT technologies. Diverse methods have been utilized for crop yield prediction or forecasting in digital agriculture. In this paper, we use environment and yield data collected in paprika smart farms. Firstly we build an optimal model relating the environment and yield. Then based on the optimal model, a methodology for yield prediction is presented. In experiment, we achieved $R^2 = [0.9171, 0.9714]$ that indicates the feasibility of predicted yields. We expect that this result will be utilized in decision making systems.

1. 서론

다양한 IoT 환경센서에 기반하여 작물 생산량을 예측하는 것은 최근 많은 관심을 받고 있다. 실제 스마트팜 경영주는 매주 온도, 습도, CO2 등의 환경을 조절하면서 원하는 생산량을 얻을 수 있도록 노력한다.

이 과정에서 예측된 미래 생산량을 자동으로 알려주면 경영주에게 큰 도움이 될 것이다.

선형회귀, SVM 등의 기계학습에서 최근에는 딥러닝을 활용하는 예측 연구가 꾸준히 진행되고 있다¹⁻⁷⁾. 생산량 예측(yield prediction)을 위해 다양한 방법들이 제안되었다. 주로 딥러닝 기반으로 구현하는데,

* This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2021-2018-0-01433) supervised by the IITP (Institute for Information & Communications Technology Promotion)

• First Author : Kangwon National University, Dept. of Electrical & Electronics Eng., ihwang@kangwon.ac.kr, 정희원

^o Corresponding Author : Kangwon National University, Dept. of Computer Eng., manbae@kangwon.ac.kr, 정희원

* Gangwon Agriculture Research & Extension Services (ARES), nhs1002@korea.kr

** Kangwon National University, University-Industry Cooperation Foundation, saneya@kangwon.ac.kr

논문번호 : 202109-232-0-SE, Received July 30, 2021; Revised September 28, 2021; Accepted October 5, 2021

Alhnaity 등은 RNN-LSTM으로 작물 생산량을 예측한다^[1]. Elavarasan 등은 강화학습으로 작물 생산량 예측 방법을 제안하였다^[2]. 선형회귀 또한 생산량 예측에 사용되고 있는데, Nagini 등은 선형회귀를 이용하여 미래 생산량을 예측하였다^[3]. Na 등은 환경과 생산량의 연간 데이터를 이용하여, 생산량과 환경의 최적 선형모델을 조사하였다^[4]. Choi 등은 패널회귀모형을 이용하여 마늘의 연간 생산량을 추정하였다. 이 과정에서 유의한 환경변수도 조사하였다^[5].

상기 딥러닝 또는 기계학습 기반 방법은 현재에서 예측할 수 있는 기간이 다음 주 또는 길면 3~4주 정도의 단기 예측이다. 즉 현재주차에서 먼 미래주차때의 생산량을 예측하면 정확도가 크게 낮아지는 단점이 있다.

따라서, 이 문제를 해결하기 위해 본 연구에서는 다중회귀분석(multiple regression analysis)기반으로 생산량 예측 방법을 제안한다. 다중회귀분석으로 얻어진 최적모델을 기반으로 생산량 예측을 수행함으로써 기존의 단기 예측의 극복이 가능하다.

생산량 예측과 달리 환경과 생산량의 최적 모델(optimal model)도 오랫동안 연구되어 왔다^[4]. 기존 연구들은 최적모델과 생산량 예측을 독립적으로 분리되어 별개로 구분하였다. 이외는 달리, 본 논문에서는 먼저 최적모델을 구축하고, 이 최적모델을 활용하여 생산량을 예측하는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 다음 장에서는 분석자에 대해 살펴보고, 제안방법을 소개한다. 3장에서는 생산량과 환경데이터의 최적 모델을 구축하는 방법을 설명하고, 이어서 4장에서는 3장의 최적모델을 기반으로 생산량 예측방법을 소개한다. 5장에서는 실험결과를 분석하고, 결론은 6장에서 정리한다.

II. 데이터 취득 및 제안 방법

분석에 사용된 자료는 강원도에 소재하는 파프리카 스마트팜 농가로부터 수집된 파프리카의 생산량과 환경 데이터이다. 환경 및 생산량 데이터는 강원도농업기술원으로부터 제공받았다. 연도는 2019년이다.

대상 온실 스마트팜에서는 환경 IoT 기기로부터 매 시간마다 환경상태가 수집된다. 이렇게 측정된 환경정보를 모형의 설명변수로 활용하여 생산량과의 연관성을 조사한다. 생산량 예측 모형의 반응변수로 사용된 생산량은 일주일간 누적된 수확량이다.

일반적으로 농작물이 일주일 간의 환경상태에 반응하는 것을 이용하여 환경변수들을 1주일 단위로 변환

표 1. 8 종류의 환경데이터
Table 1. Eight environment data

Environment	Unit
Inside Temperature (IT)	°C
Outside Temperature (OT)	°C
Inside Humidity(IH)	%
Accumulated Light	Watt/m ²
EC	mS/cm
Light	Watt/m ²
CO2	ppm
DIF	

한다. 환경정보는 내부온도, 내부습도, 일사량 등으로 구성되는데, 환경데이터는 1시간마다 측정되기 때문에, 1주일의 평균값을 이용한다.

환경정보는 내부온도, 외부습도, 내부습도, 누적일사량, EC, 광량, CO2, 및 낮과 밤의 온도차(DIF) 등으로 분류되며, 수집한 환경데이터에 대한 자세한 설명은 다음의 표 1과 같다.

그림 1은 파프리카 볼란테 농장에서 수확한 주간 생산량을 보여준다. 단위는 Kg이다. 조사일은 17주차인 4월 28일부터 1주마다 생산량을 측정하였고, 12월 16일이 종료일이다. 그래프에서 보듯이, 주차별 생산량의 큰 변화는 분석에 어려움을 준다. 이 변화는 일반적인 현상으로 수확 작물량은 주차별 일정하지는 않다. 기계학습으로 이 문제를 극복하는 것이 필요하다.

그림 2는 생산량 예측의 전체 흐름도를 보여준다. 먼저 회귀분석을 기반으로 8개의 환경 데이터와 생산량 데이터의 최적 모델을 구축한다. 다음에는 이 최적

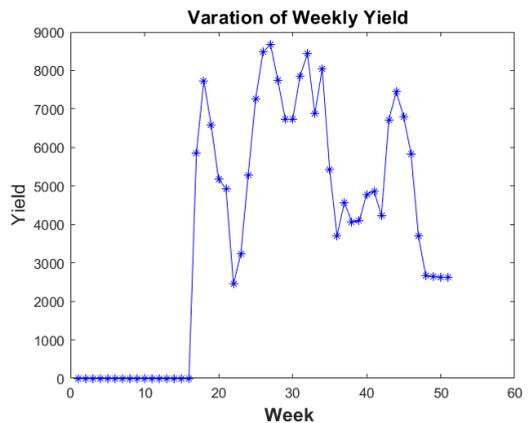


그림 1. 주차별 생산량의 변화 (단위: kg)
Fig. 1. Variation of weekly yield (unit: kg)

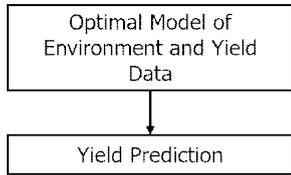


그림 2. 생산량 예측의 전체 흐름도
Fig. 2. Overall flow diagram of yield prediction

모델을 토대로 생산량 예측을 하게된다. 최적 모델은 3장에서 소개하고, 최적모델을 기반으로 생산량 예측은 4장에서 자세하게 설명된다.

III. 최적 모델 구축

생산량은 $y = [y_{P1}, \dots, y_{P2}]$, 환경인자는 $e = [e_{T1}, \dots, e_{T2}]$ 의 값을 가지고 있다. $P1$ 은 생산량 시작 주차이고, $P2$ 는 생산량의 마지막 주차이다. 마찬가지로, $T1$ 은 환경센서의 시작 주차이고, $T2$ 는 마지막 주차이다. 일반적으로 $P1 > T1$, $P2 > T2$ 이다.

그림 3은 생산량과 환경의 모델 구축의 전체 흐름도를 보여준다. 생산량, 환경데이터가 주어지면 상관계수(correlation coefficient) 및 대응주차(corresponding week)를 계산한다. 대응주차는 일정구간의 생산량과 환경의 전체 기간 $[T1, T2]$ 에서 일정구간의 환경데이터의 상관계수를 구한 후에, 최대 상관계수를 주는 환경데이터의 주차를 의미한다. 다중선형회귀를 이용하여 회귀 파라미터 및 대응주차값을 계산한다.

주차마다 D 구간(interval)의 생산량 데이터 y 와 환경데이터 e 가 주어진다. D 는 보통 8~10주이다. 현재 주차가 w_c 이면 생산량 y 는 다음과 같다.

$$y = [y_{w_c-(D-1)}, \dots, y_{w_c-1}, y_{w_c}] \quad (1)$$

환경데이터 e 는 w_c 의 이전 주차에서 모두 얻어진다.

$$\vec{e} = [e_{w_p-(D-1)}, \dots, e_{w_p-1}, e_{w_p}], \quad (2)$$

여기서 $w_p = \{D, D+1, \dots, w_c - 1\}$ 이다.

생산량 데이터는 $D \times 1$ 벡터이고, 환경데이터는 M 개의 환경정보가 있으므로, $D \times M$ 행렬로 표현된다. 표 1에서 보는 것처럼, $M=8$ 이다.

3.1 상관계수 및 대응주차 계산

M 개의 환경인자가 있고, 각각의 환경인자는 다른 주차에서 생산량에 영향을 주게 된다. p 주차에서의 생산량은 t 주차의 환경과 연관성이 높다고 하면, t 주차는 $p (> t)$ 주차의 대응주차이다. 대응주차는 환경인자마다 다른 값을 가질 수 있다. 즉, 각각의 환경인자는 다른 주차에서 영향을 주게 된다. 이 영향력을 조사하기 위해서 통계분석 기술 중에서 상관도 지수인 상관계수를 활용한다.

생산량 주차 p ($p \in \{P1, P2\}$)에서 D 구간의 생산량 $y_p = [y_{p-(D-1)}, \dots, y_{p-1}, y_p]$ 을 구성하고, 환경주차 t ($t \in \{T1, T2\}$)에서 D 구간의 환경데이터 $e_t = [e_{t-(D-1)}, \dots, e_{t-1}, e_t]$ 을 구성한다. 다음에 y_p 와 e_t 의 상관계수 ρ 를 구한다. y_p 와 e_t 를 다음과 같이 간단히 표현하면

$$\begin{aligned} y_p &= [y_1, y_2, \dots, y_D], \\ e_t &= [e_1, e_2, \dots, e_D], \end{aligned} \quad (3)$$

두 값의 상관계수 ρ 는 다음과 같이 구해진다.

$$\rho = \frac{\sum_{i=1}^D (e_i - \bar{e})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^D (e_i - \bar{e})^2} \sqrt{\sum_{i=1}^D (y_i - \bar{y})^2}} \quad (4)$$

여기서 \bar{e} 는 e_t 의 평균값이고, \bar{y} 는 y_p 의 평균값이다. 생산량 주차 p 의 범위는 $[P1, P2]$, 환경 주차 t 의 범위는 $[T1, T2]$ 이다. 따라서 p 에서 얻을 수 있는 주차의 개수는 $S = P2 - P1 - D$ 개이다. 유사하게, t 에서 얻을 수 있는 주차 개수는 $R = T2 - T1 - D$ 개이다. 각 p 에서의 생산량 데이터와 R 개의 환경 데이터와의 상관계수를 구한 다음에, 최대 상관계수를 가지

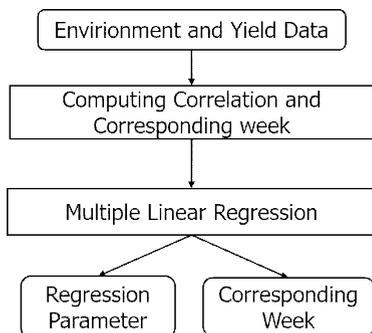


그림 3. 환경-생산량 최적 모델의 흐름도
Fig. 3. The flow diagram for optimal model of environment and yield

는 환경 대응주차 τ 를 구한다.

M 개의 환경인자가 있고, 주차 p 의 생산량은 y_p 인데, 첫 번째 환경인자 e_1 는 대응주차 τ_1 , 두 번째 e_2 는 τ_2 , 마지막 e_M 은 τ_M 이라고 가정하면, 이 대응주차에서 각 환경과 생산량의 상관도가 가장 크다고 할 수 있다.

3.2 다중 선형회귀 분석

다중회귀모델은 다음과 같이 정의된다.

$$Y = X\beta \tag{5}$$

다중회귀 분석은 종속변수인 생산량 변수 Y 를 설명하는데, M 개의 독립변수인 x_1, x_2, \dots, x_M 를 도입한다.

종속변수 Y 는 D 구간의 생산량인데, 특정 주차 p 에서 이전 $(p-D+1)$ 까지 생산량이다.

$$Y = \begin{bmatrix} y_{p-D+1} \\ y_{p-D+2} \\ \vdots \\ y_p \end{bmatrix} \tag{6}$$

X 는 M 개의 환경인자이고, 각 열은 D 개의 환경 데이터이다. 각 데이터는 대응주차에 있는 값들이다.

$$X = \begin{bmatrix} 1 & e_1 & e_2 & \dots & e_M \\ 1 & e_1 & e_2 & \dots & e_M \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & e_1 & e_2 & \dots & e_M \end{bmatrix} \tag{7}$$

회귀 파라미터는

$$\beta = [\beta_0, \beta_1, \dots, \beta_D]^T \tag{8}$$

이다. T 는 전치행렬이다.

해 β 를 구하기 위해서 다음과 같이 변환하고

$$(X^T X)\hat{\beta} = X^T Y \tag{9}$$

최종적으로 다음식에서 β 값이 결정된다.

$$\hat{\beta} = (X^T X)^{-1} X^T Y \tag{10}$$

여기서 $\hat{\beta}$ 는 회귀 파라미터이고, β_0 는 절편(offset) 값이다.

각 생산량의 주차별로 선형방정식이 얻어지고, 회귀 파라미터와 대응주차가 저장된다.

3.3 최적 모델 생성

얻어진 선형회귀식은 표 2에서 보여준다. 17주차~45주차까지의 회귀식 중, 17, 24, 30, 35 주차를 보여준다. 식에서 ()안의 숫자는 대응주차이다. 17주차에서 생산량은 다음 식으로 표현된다.

$$\begin{aligned} \text{Yield} = & 1.8399 + 11.8399 \cdot \text{EC}(13) - \\ & 3.615 \cdot \text{외부온도}(16) \\ & - 0.077711 \cdot \text{CO}_2(16) - 0.0031175 \cdot \text{DIF}(15) \end{aligned} \tag{11}$$

식 (11)을 보면 이 주차의 생산량은 13주차때의 EC, 16주차의 외부온도, 16주차의 CO₂, 그리고 15주차의 DIF와 연관성이 있다는 것으로 보여준다.

표 2. 환경데이터와 생산량과의 관계를 보여주는 주차별 선형회귀식. ()는 대응주차임
Table 2. Weekly linear regression formula showing the relationship between environment and yield. () is a corresponding week

Week	Linear regression formula
17	Yield = 1.8399 + 11.8399 • EC(13) - 3.615 • OT(16) - 0.077711 • CO ₂ (16) - 0.0031175 • DIF(15)
24	Yield = -18.8688 - 18.8688 • Acc_light(19) - 0.0084419 • EC(16) + 2.1548 • Light(19) + 0.049211 • OT(20) + 0.18682 • CO ₂ (23) + 0.017905 • DIF(20)
30	Yield = 17.5415 + 17.5415 • IH(27) - 0.091524 • CO ₂ (25) - 0.02691 • DIF(24)
35	Yield = 68.0638 + 68.0638 • IT(28) - 3.4883 • OT(26) - 0.024688 • CO ₂ (27) + 19.7567 • DIF(27)

IV. 생산량 예측

생산량 예측은 현재주차까지의 환경데이터를 분석하여 미래의 생산량을 예측하는 것이다. 그림 4는 주차의 관계를 보여주는데, 현재주차(current week)가 w_c 이면, 과거주차(past week)는 w_p , 미래주차(future week)는 w_f 이다. 과거주차 w_p 기간의 환경데이터를 기반으로 생산량을 예측한다. 최적모델과 생산량예측을 독립적으로 진행하지 않고, 3장의 최적모델을 기반으로 생산량 예측 방법을 제안한다.

그림 5는 생산량 예측의 블록도를 보여준다. 그림

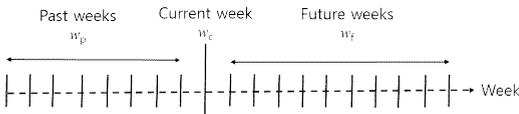


그림 4. 현재주차, 과거주차, 미래주차의 관계
Fig. 4. Relation of current, past, future weeks

4에서 보는 것처럼, 미래 생산량 예측은 현재 주차 w_c 에서 미래 주차 w_f 의 생산량을 예측하는 것이다. $w_f = [w_c + 1, w_c + 2, \dots, w_c + J]$ 로 J 개의 미래주차가 있다. 미래 생산량 예측에는 w_c 와 w_p 의 환경데이터가 필요하다. 과거주차 $w_p = [1, \dots, w_c - 2, w_c - 1]$ 로 $J = w_c - 1$ 개가 있다.

그림 6은 예측 생산량을 저장하는 PROD와 누적 개수를 저장하는 ACC의 구조를 보여준다. 출력으로 예측된 생산량을 저장하는 I 개의 빈(bin)으로 구성된 PROD, 같은 크기의 누적기(accumulator) ACC가 있다. 초기에는 모든 빈 값은 0이다.

특정 현재주차 w_c 에서는 w_c 및 J 개의 과거주차 환경데이터가 있다. w_c 는 1개이므로 총 $(I+1)$ 개의 환경데이터가 준비된다.

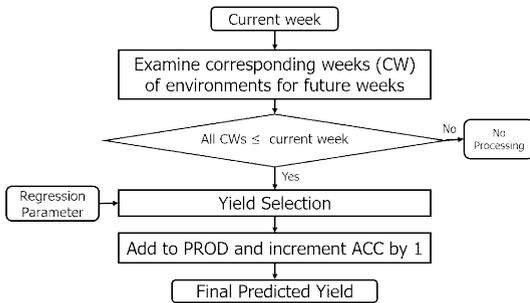


그림 5. 생산량 예측 블록도
Fig. 5. The block diagram of yield prediction

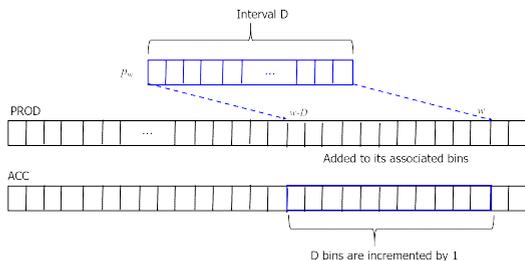


그림 6. 예측 생산량은 PROD에 더해지고, ACC 빈 값을 1 증가함
Fig. 6. Predicted yield is added to PROD and ACC bin is incremented by one.

I 개의 미래주차 w_f 의 생산량을 구하기 위해서, 표 2를 이용하여 각 w_f 의 생산량과 대응주차를 조사한다. 이 대응주차 값이 한 개라도 w_c 보다 크면 현재 주차에 대한 생산량 예측은 실행하지 않는다. 반대로 이 조건을 만족하면, 생산량을 PROD에 더한다. 주차별 생산량은 D 구간을 이용하기 때문에 미래주차가 w_f 이면, $[w_f - D, w_f]$ 의 구간에 생산량이 더해진다. 즉, PROD의 해당 D 개의 빈(bin)에 더해진다. 또한 ACC의 D 개의 빈 값을 1 증가한다.

미래주차의 예측생산량이 모두 더해지면, 최종적으로 평균값을 구하여 예측된 생산량 Y_w 을 얻는다

$$Y_{w_f} = \frac{\text{PROD}_{w_f}}{\text{ACC}_{w_f}} \quad (12)$$

V. 실험 결과

분석에 사용된 자료는 강원도에 소재하는 파프리카 스프라트팜 농가로부터 수집된 2019년도 파프리카의 생산량과 환경 데이터이다. 환경 및 생산량 데이터는 강원도농업기술원으로부터 제공받았다. 데이터에 대한 내용은 2장에서 자세히 기술하였다.

그림 7은 제안방법으로 얻은 생산량 예측값을 보여준다. 가로축은 주차이고, 세로축은 생산량이다. 청색 선은 그림 1의 실제값이고, 적색선은 예측 생산량 값인 식 (13)의 Y_{w_f} 이다. 예측을 하는 현재주차는 17, 20, 25, 43이다. 오차는 있지만 예측값이 실제값을 따라가는 것을 관측할 수 있다. 결과를 봤을 때 오차가 매우 작다고 보기는 어렵다고 판단되는데, 근본 원인은 주차별 생산량의 급격한 변화이다. 이를 해결하기 위해서는 생산량의 이동 평균을 구하면 오차는 감소하게 된다. 그렇지만 본 실험에서는 이동 평균을 사용하지 않는다.

그림 7(a)를 보면, 1주차~17주차까지의 환경데이터로부터 18주차~29주차까지의 생산량을 예측한다. 적색 예측값이 0인 것은 예측이 불가능한 경우이다. 이유는 예를 들어 32주이면 $[32 - D, 32]$ 주차의 생산량들은 구간을 만족하는 환경의 대응주차가 없기 때문이다. D 는 회귀 분석 구간이다.

25주차에서 예측한 결과는 그림 7(b)에서 보여진다. 26주차~39주차의 예측 생산량을 얻는다. 그림 7(c)는 35주차에서 예측한 결과를 보여준다. 36주차~49주차까지의 생산량을 얻는다. 43주차에서 예측

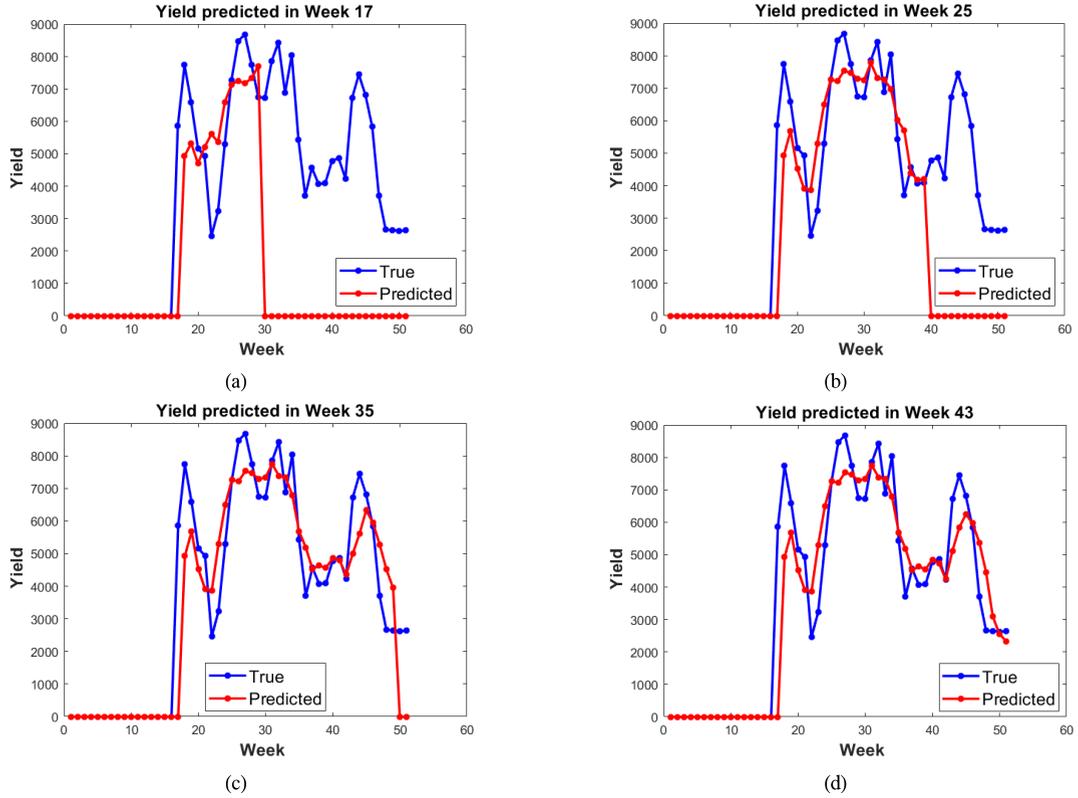


그림 7. 환경-생산량 최적모델로 얻은 생산량 예측값. (a) 17주차, (b) 25주차, (c) 35주차, 및 (d) 43주차
 Fig. 7. Yields predicted in (a) Week 17, (b) Week 25, (cc) Week 35, and (d) Week 43

한 결과는 그림 7(d)에서 보여진다. 44주차~52주차의 예측 생산량을 얻는다.

예측 데이터의 적합도를 평가하기 위해서 결정계수인 R^2 를 이용한다. 그림 8은 17주차~45주차 기간의 R^2 의 변화를 보여준다. R^2 값은 [0.9171, 0.9714]에서 분포되어 있기 때문에 제안방법의 예측성능은 신

뢰성이 높다.

VI. 결 론

생산량 예측은 스마트팜을 운영하는 경영주에게 중요한 정보이다. 이를 기반으로 의사결정시스템의 구축이 가능하기 때문이다. 본 논문에서는 다항회귀분석을 이용하여 환경 대응주차 및 회귀 파라미터를 얻은 후에, 환경-생산량의 최적모델을 구축하였다. 이를 기반으로 생산량을 누적한 후, 평균값으로 예측 생산량을 결정하는 방법을 제안하였다. 파프리카 스마트팜 데이터를 이용한 실험에서는 17주차~45주차 기간 동안의 예측 생산량 방법의 적합도를 조사한 결과, R^2 값은 [0.9171, 0.9714]의 분포를 가져, 신뢰성 있는 예측 성능을 얻었다. 미래 생산량예측 방법은 향후 의사결정 모델에 활용할 계획이다.

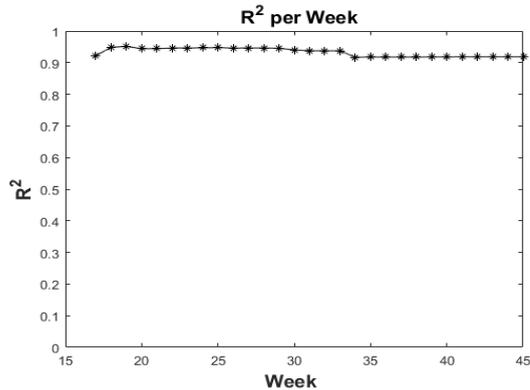


그림 8. 주차별 R^2 값
 Fig. 8. R^2 values per week

References

- [1] B. Alhnaity, S. Pearson, G. Leontidis, and S. Kollias, "Using deep learning to predict plant growth and yield in greenhouse environments," *arXiv:1907.00624v1* [cs.LG], Jul. 2019.
- [2] D. Elavarasan and P. Durairajvincent, "Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications," *IEEE Access*, Digital Object Identifier 10.1109/ACCESS.2020.2992480, May 2020.
- [3] S. Nagini, T. Kanth, and B. Kiranmayee, "Agriculture yield prediction using predictive analytic techniques," *2nd Int. Conf. Contemporary Computing and Informatics (IC3I)*, Dec. 2016.
- [4] M. Na, Y. Park, and W. Cho, "Study on optimal environmental factors for tomato using smart farm," *J. Korean Data & Info. Sci. Soc.*, vol. 28, no. 6, 2017.
- [5] S. Choi and J. Baek, "Garlic yields estimation using climate data," *J. Korean Data & Info. Sci. Soc.*, vol. 27, pp. 969-977, 2016.
- [6] P. Song and J. Kim, "Correlation and regression analyses based on variable selection for the university evaluation index," *J. Korean Data & Info. Sci. Soc.*, vol. 23, pp. 457-465, 2012.
- [7] P. Surya and I. Aroquiaraj, "Crop yield prediction in agriculture using data mining predictive analytic techniques," *IJRAR*, vol. 5, no. 4, Dec. 2018.

황 인 철 (In-Chul Hwang)



1993년 2월 : 고려대학교 전자공학과 졸업
 1995년 2월 : 고려대학교 전자공학과 석사
 2000년 2월 : 고려대학교 전자공학과 박사
 2001년 3월~2007년 2월 : 삼성 전자 책임연구원

2007년 3월~현재 : 강원대학교 전기전자공학과 교수
 <관심분야> 저전력 SoC설계, 지능형반도체설계, 머신러닝

[ORCID:0000-0002-4702-8276]

노 희 선 (Heesun Noh)



1995년 2월 : 서울시립대학교 환경예학과 졸업
 1997년 2월 : 서울시립대학교 환경예학 석사
 2001년 2월 : 서울시립대학교 환경예학 박사
 1996년 8월~현재 : 강원도농업기술원 농업연구사

<관심분야> 디지털농업, 농업빅데이터 분석, 생산성 향상 모델 개발, 스마트팜 농작업 표준모델 개발

[ORCID:0000-000-4702-8276]

양 동 일 (Dong-II Yang)



2002년 2월 : 강원대학교(삼척캠퍼스) 컴퓨터공학과 졸업
 2004년 2월 : 강원대학교 컴퓨터과학과 이학석사
 2007년 8월 : 강원대학교 컴퓨터과학과 이학박사
 2013년 1월~2019년 7월 (주)오픈시스템서비스 부장

2019년 8월~현재 : 강원대학교 산학협력교수
 <관심분야> 데이터 분석, 온톨로지, 소프트웨어공학

[ORCID:0000-0003-0169-8853]

김 만 배 (Manbae Kim)



1983년 2월 : 한양대학교 전자공
학과 졸업

1986년 2월 : University of
Washington 전기공학과 석사

1991년 2월 : University of
Washington 전기공학과 박사

1991년 3월~1997년 2월 : 삼성
종합기술원 수석연구원

1997년 3월~현재 : 강원대학교 컴퓨터정보통신공학과
교수

<관심분야> 영상처리, 컴퓨터비전, 데이터 분석

[ORCID:0000-0002-4702-8276]