

01

데이터 기반 작물 생육 모의와 수확량 예측 알고리즘



김 마 가
서울대학교
지역시스템공학전공
/ 박사과정
mnkm53@snu.ac.kr



최 진 용
서울대학교
생태조경·지역시스템공학부
/ 교수
iamchoi@snu.ac.kr

1. 머리말

최근 4차 산업혁명의 발전에 따라 빅데이터, 인공지능, IoT 등을 접목한 새로운 형태의 산업이 다양한 분야에서 등장하고 있다. 4차 산업은 데이터를 기반으로 물리 세계와 디지털 세계를 융합하는 것을 특징으로 하며, 방대한 자료로부터 새로운 정보를 도출하고 실시간 통신과 추론, 예측 등을 통해 기존과는 다른 형태의 서비스와 산업 구조를 창출하고 있다. 스마트 제조, 자율주행, 디지털 헬스케어 등이 대표적인 사례에 해당한다(World Economic Forum, 2016).

이러한 흐름에 따라 농업 분야에서도 빅데이터, 인공지능, IoT 기술을 접목한 스마트 농업이 빠르게 확산되고 있다. 스마트 농업은 센서와 데이터, 정보통신기술을 활용하여 작물 재배 환경을 정밀하게 관리하는 농업으로, 기존의 경험에 의존한 관행적 농업에서 벗어나 데이터 기반의 의사결정을 수행하는 것이 특징이다.

스마트 농업의 핵심은 다양한 센서를 활용한 데이터 수집에 있다. 이러한 시스템은 기상 정보뿐만 아니라 토양 환경, 작물 생육 상태, 생육 스트레스, 재배 환경 등 여러 정보를 실시간으로 수집·분석함으로써, 정밀 농업을 구현하고 한정된 자원을 효율적으로 활용할 수 있도록 한다. 이를 통해 작물의 생육 및 스트레스를 효과적으로 제어하고, 안정적이면서도 높은 생산성과 품질을 확보하는 것을 목표로 한다. 스마트 농업에서 수집·활용하는 데이터는 기상 자료, 토양 수분 및 온도, 양분 상태, 작물 생육 지표, 영상 기반 생육 정보 등으로 구성되며, 수집한 데이터는 생육 모니터링, 병해 예측, 수확량 예측 등에 활용되고 있다.

본 고에서는 스마트 농업의 데이터 활용 사례 중 하나로, 작물 생육 상태

를 기반으로 하는 수확량 예측 알고리즘을 소개하고, 그 적용 가능성과 시사점을 논하고자 한다.

2. 연구배경 및 개요

본 고에서 소개하는 작물 수확량 예측 알고리즘은 농업생산 시범단지의 지능형 플랫폼 구축을 위해 개발되었으나, 개발 당시 시범단지 조성이 초기 단계에 있었기 때문에 데이터 수집 기간이 짧고, 확보된 자료 또한 제한적인 상황이었다.

특히 작물은 재배 기간이 길고 수확 직전까지 정확한 생산량을 파악하기 어려우며, 조사 과정 또한 노동집약적이다. 이로 인해 작물 수확량 자료는 단기간에 충분히 확보하기 어려운 특성을 가지며, 기상, 토양, 생육 상태와 같은 다른 데이터에 비해 수집이 제한적이다. 이에 더해 농업은 기상 조건의 영향을 크게 받기 때문에 평균적인 기상 조건뿐만 아니라 일 단위의 극한 기상이나 생육 시기별 기상 조건에 따라서도 수확량이 크게 달라진다. 이러한 특성으로 인해 다양한 조건에서 재배된 실제 수확량 자료를 확보하는 데에는 현실적인 제약이 따른다.

이와 같은 이유로 스마트 농업단지의 초기 단계에서는 데이터 기반의 수확량 예측 알고리즘을 구축하는 데 어려움이 존재한다. 이에 본 연구에서는 작물 생육 모형을 활용하여 다양한 기상 조건에서 작물 생육을 모의하고, 이를 기반으로 작물 수확량을 예측하는 알고리즘을 개발하고자 하였다.

3. 작물 생육 모형 개요 및 기상 시나리오 구축

본 연구에서는 제한된 시간과 인력 여건으로 인

해 시험재배를 수행하기 어렵고, 스마트 농업단지 초기 단계에서 축적된 자료가 부족한 점을 고려하여 작물 생육 모형인 APEX 모형을 활용하였다. 이를 통해 콩, 벼 (조생종, 중만생종), 양파, 마늘에 대한 수확량 예측 알고리즘을 개발하였으며, 본 고에서는 콩과 벼를 중심으로 알고리즘 개발 과정과 결과를 고찰하고자 한다.

APEX 모형은 필지 및 소유역 단위에서 농업 활동과 토양·수문·양분 이동을 통합적으로 모의할 수 있는 모형으로, 기상, 토양, 재배관리 조건을 반영하여 작물 생육과 수확량을 모의할 수 있다. 작물 생육 모의 과정에서는 작물의 생육 매개변수를 기반으로 기상 조건과 토양 상태, 재배관리 정보를 종합적으로 반영하여 생육 과정을 재현한다. 이 과정에서 복사에너지, 엽면적지수, 기상 조건 등을 통해 잠재 성장률을 산정하고, 이후 수분, 온도, 양분 스트레스 요인 등을 고려하여 최종 수확량을 산출한다. 그림 1은 APEX 모형 입출력 자료의 구조도를 나타낸 것이다(Steglich et al., 2023; Williams et al., 2023).

본 연구에서는 농촌진흥청에서 제공하는 콩과 벼의 국내 작물 모수(최순균 외, 2017; 허승오 외, 2019)를 활용하여 우리나라 환경에 적합한 생육 모의를 수행하였다. 또한 APEX 모형은 파종 시기나 재배 일정과 같은 영농 조건을 유연하게 조정할 수 있어 다양한 환경 조건에서 수확량 모의가 가능하며, 본 연구에서는 파종 시기의 변화를 고려하여 재배관리 시나리오를 구성하고, 이를 모의에 적용하였다.

APEX 모형의 작물 생육 모의를 위해 작물 모수 외에 기상, 토양, 재배관리 자료를 체계적으로 구축해야 한다. 특히 본 연구에서는 다양한 기상 조건을 반영하기 위해 과거에 관측된 기상자료를 기반으로 연 단위 기상 시나리오를 구성하였다.

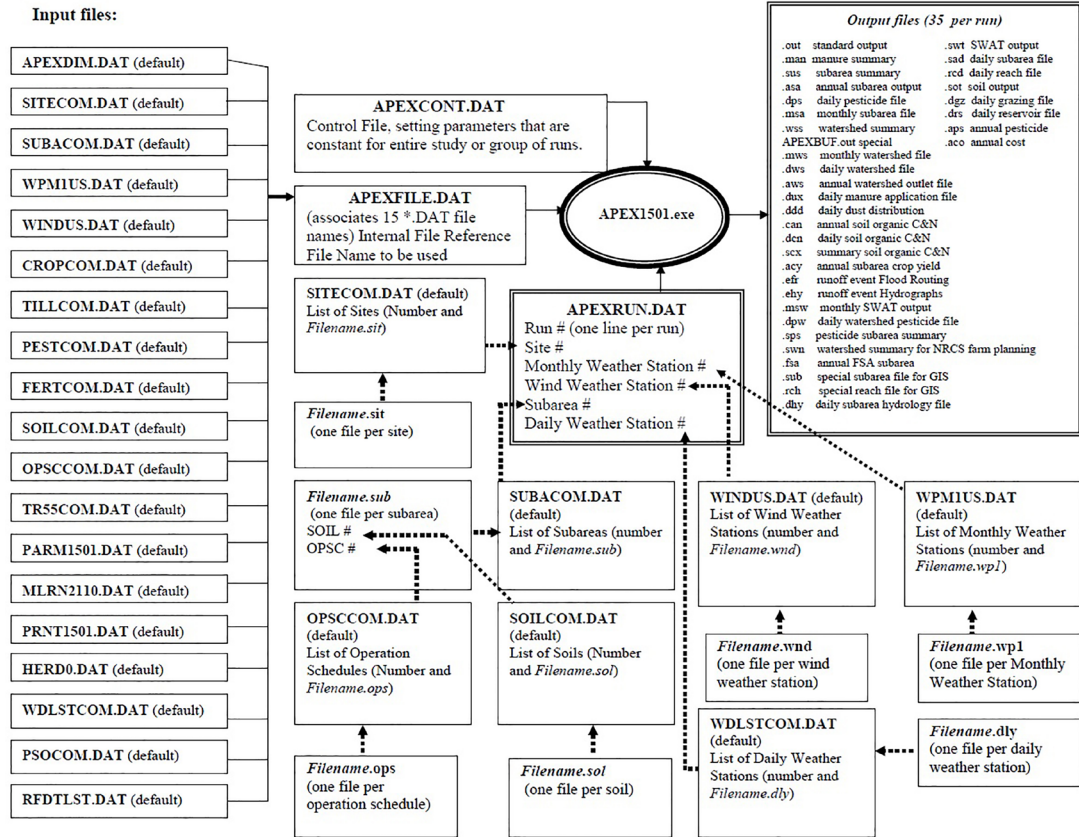


그림 1. APEX 모형 입출력 자료 구조도 (Steglich et al., 2023)

표 1. 대상지구 인근 종관기상관측소 자료 구성 및 위치도

관측소번호	관측소명	자료축적기간 [년]	시나리오 수	관측소 위치도
156	광주	1940-2021	82	
165	목포	1906-2021	116	
172	고창	2011-2021	11	
174	순천	2012-2021	10	
247	남원	1973-2021	49	
251	고창군	2008-2021	14	
252	영광군	2008-2021	14	
254	순창군	2009-2021	13	
256	주암	1973-2021	39	
258	보성군	2011-2021	11	
259	강진군	2010-2021	12	
260	장흥	1973-2021	49	

이를 위해 대상 지역인 나주 인근 12개 종관기상 관측소의 2021년 까지 장기 기상자료를 수집하였으며, 이를 인근 지역에서 실제 발생 가능한 기상 시나리오로 간주하여 총 420개의 연 기상 시나리오를 구축하고 모형에 적용하였다. 표 1에 기상 시나리오 정보를 정리하고, 관측소의 위치와 분포를 도시하였다.

4. APEX 모형 입력자료 구축

APEX 모형을 활용한 작물 생육 모의를 위해 기상 자료 외에 토양 및 재배관리 자료의 구축이 필요하다. 본 연구에서는 토양 자료로 농촌진흥청에서 제공한 APEX 토양통 자료를 활용하였으며, 농촌진흥청 흙토람 (<https://soil.rda.go.kr>) 토양통계 자료를 기반으로 작물별 최적 토양통을 선정하여 적용하였다.

재배관리 자료의 경우 작부체계, 시비량, 유효 적산온도, 재식밀도 등의 정보를 기반으로 구성하였으며, 농업기술길잡이(농촌진흥청, 2014; 농촌진흥청, 2021), 농촌진흥청 농사로(<https://www.nongsaro.go.kr>)와 농업기술센터 자료, 연구 문헌(윤푸른, 2022) 등을 참고하여 재배관리 기초자료를 수집하였다. 유효적산온도와 재배기간의 경우 대상지역에서 시험 재배를 수행한 전문

가 자문을 통해 실제 재배 조건을 반영할 수 있도록 하였다.

파종 및 이앙 시기의 변화를 고려하기 위해 농사로와 농업기술센터에서 수집한 파종 및 이앙 시기를 5일 간격으로 구분하여 다양한 재배관리 시나리오를 구성하였다. 표 2는 작물 재배관리 기초자료와 파종 및 이앙 시기에 따른 시나리오를 정리한 것이다.

5. 작물 생육 모형 결과 검토 및 생육 상태 기반 수확량 예측 알고리즘

작물의 생육은 장기간에 걸쳐 기상, 토양, 재배관리 등 다양한 환경 요인의 영향을 받으며 진행되기 때문에, 이를 단일 값으로 환원하여 수확량을 예측하는 데에는 한계가 있다. 이러한 점을 고려하여 본 연구에서는 구축한 기상 시나리오를 토대로 작물 생육 모의를 수행하고 시나리오에 따른 파종 및 이앙 시기별 수확량을 단일 값이 아닌 분포 형태로 제시하고자 하였다. 그림 2는 파종 및 이앙 시기별 수확량 모의 결과 분포 그래프를 도시한 것이며, 전체 결과는 표 3과 표 4에 정리하였다. 본 연구에서 적용한 조건 하의 모의 결과에 따르면 수확량이 최대가 되는 콩의 최적 파종시기는 5월 16-26일(case 4, 5, 6), 조생종벼의 최

표 2. 작물 재배관리 기초자료와 파종 및 이앙 시기에 따른 시나리오

작물	재배기간 [days]	표준시비량 (N-P-K) [kg/10a]	적산온도 [°C]	파종시기	파종 시나리오
콩	130	3.0-1.31-2.82	2,750	5월 상순 - 6월 하순	5/1, 5/6, ..., 6/30
벼 (조생종)	120	9.0-1.96-4.73	1,824.21	5월 중순 - 6월 중순	5/11, 5/16, ..., 6/20
벼 (중만생종)	140	9.0-1.96-4.73	2,071.63	5월 상순 - 6월 중순	5/1, 5/6, ..., 6/20
	145		2,088.41		
	149		2,099.39		

적 이앙시기 5월 26일-6월10일(case 5, 6, 7)에 해당하는 것으로 나타났으며, 해당 시기의 평균 수확량은 생중량 기준 콩 1.51 ton/ha, 조생중벼 6.26 ton/ha로 모의되었다. 이는 단일 작물 모수를 기반으로 하는 모형 모의 결과로, 새로운 작물 모수나 환경조건을 적용할 경우 결과에 차이가 발생할 수 있다. 다만 농촌진흥청 농사로(<https://www.nongsaro.go.kr>)의 '이달의 농업기술'에서 제시된 바와 같이 파종 시기가 이르거나 늦을 경우 수량이 감소하는 경향이 나타나, 본 연구의 결과가 파종 및 이앙 시기에 따른 수량 변화를 일정 부분 반영하고 있는 것으로 보인다.

본 연구의 수확량 모의 결과는 실제 수확량 데이터가 충분하지 않은 상황에서 모형 모의만으로 도출된 것이므로, 타당성을 검토하기 위해 통계청(<https://kosis.kr>)의 도별 생산량 자료를 활용하여 과거 생산량과 비교 분석을 수행하였다.

1980-2021년 전라남도 대두의 평균 생산량은 1.54 ton/ha, 1982-2021년 조곡의 평균 생산량은 6.41 ton/ha로 모의된 수확량 결과와 비교해 봤을 때 콩과 벼 모두 3% 이내 차이로 나타나 작물 생육 모의에 따른 수확량 예측이 어느 정도 합리적인 범위에서 이루어졌음을 확인할 수 있다. 그림 3은 콩(대두), 벼(조곡)의 과거 생산량 분포를 도시한 것이다.

작물 수확량은 생육단계별 기상, 생육 상태에 크게 영향을 받으나, 기상 시나리오에 따른 모의만으로는 이를 반영하기 어렵다. 따라서 본 연구에서는 관측된 생육 상태에 따른 조건부 수확량 분포를 통해 수확량 예측 알고리즘에 작물 생육 상태를 반영하고자 하였으며, 이를 위해 엽면적지수와 초장을 주요 지표로 활용하였다. 특정 시점에 실제 관측된 생육 상태를 반영하기 위해 기상 시나리오별 엽면적지수와 초장의 일별 작물 생육

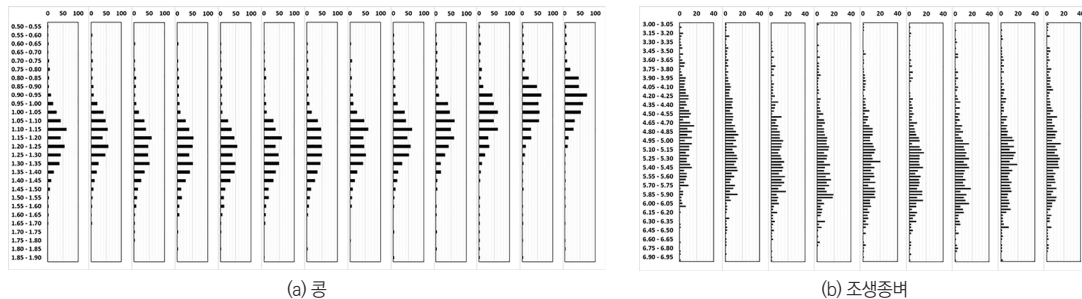


그림 2. 파종 시기별 수확량 모의 결과 분포 (건조중량 기준)

표 3. 콩 파종 시기별 수확량 모의 결과 통계 (건조중량 기준, ton/ha)

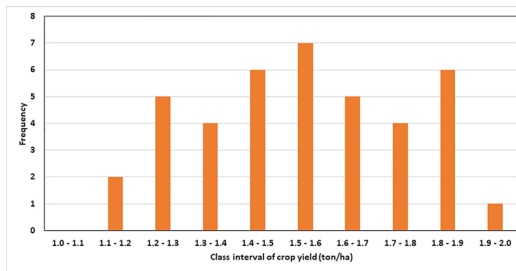
Case	1	2	3	4	5	6	7	8	9	10	11	12	13
min	0.52	0.62	0.62	0.63	0.64	0.63	0.62	0.61	0.59	0.56	0.54	0.51	0.48
max	3.05	3.11	3.11	3.11	3.16	3.12	3.01	2.85	2.80	2.69	2.56	2.36	2.25
avg	1.21	1.25	1.28	1.29	1.31	1.30	1.27	1.24	1.21	1.16	1.09	1.02	0.98
std	0.29	0.31	0.31	0.31	0.31	0.31	0.30	0.29	0.28	0.27	0.26	0.24	0.22
mdn	1.18	1.22	1.24	1.25	1.27	1.27	1.24	1.20	1.17	1.12	1.07	0.99	0.95

* min: 최소값, max: 최대값, avg: 평균, std: 표준편차, mdn: 중앙값

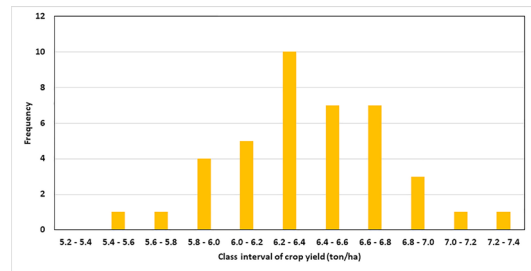
표 4. 조생종벼 이앙 시기별 수확량 모의 결과 통계 (건조중량 기준, ton/ha)

Case	1	2	3	4	5	6	7	8	9
min	2.32	2.83	2.84	2.95	3.00	3.07	2.90	2.65	2.73
max	7.65	7.82	8.10	7.96	8.02	8.38	7.98	7.43	7.05
avg	4.80	5.01	5.25	5.33	5.36	5.40	5.39	5.28	5.10
std	0.80	0.78	0.76	0.74	0.71	0.75	0.73	0.73	0.78
mdn	4.83	5.06	5.31	5.38	5.38	5.43	5.42	5.32	5.19

* min: 최소값, max: 최대값, avg: 평균, std: 표준편차, mdn: 중앙값



(a) 콩



(b) 벼

그림 3. 도별 작물 생산량 분포

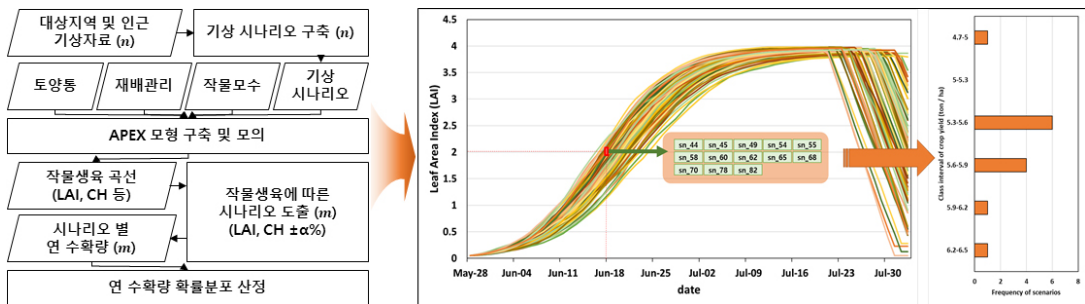


그림 4. 생육 상태 기반 수확량 예측 알고리즘 모식도 (LAI: 엽면적지수, CH: 초장)

곡선을 생성하여 관측된 자료와 유사한 생육 상태를 선별하고, 이를 통해 수확량을 예측하였다. 그림 4는 생육 상태 기반의 수확량 예측 알고리즘을 모식도로 도시한 것이다.

구체적으로는 특정 시점의 엽면적지수와 초장이 기준값 대비 일정 범위 내에 포함되는 시나리오를 선별하고, 이들 시나리오에서 도출된 조건부 수확량 분포를 이용하여 평균과 분산을 산출함으

로써 보다 현실적인 수확량 예측이 가능하도록 하였다. 그림 5는 수확량 예측 알고리즘의 실제 적용 사례로 파종시기가 5월 21일인 콩 생육 모의 결과에서 임의의 작물 생육 상태를 반영하여 수확량을 예측하는 과정을 도시한 것이다.

그림 5에서 7월 19일에 측정된 콩의 엽면적지수가 0.5, 초장이 0.7 m 라고 가정했을 때 기상 시나리오 중 엽면적지수가 $0.5 \pm 5\%$, 초장

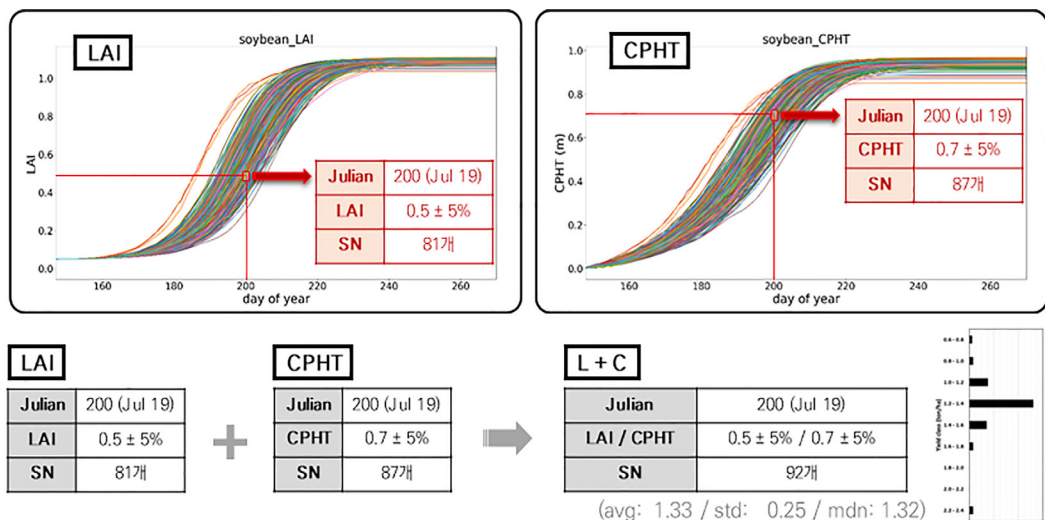


그림 5. 생육 상태 기반 수확량 예측 적용 예시 (LAI: 엽면적지수, CPHT: 초장)

이 $0.7 \text{ m} \pm 5\%$ 에 해당하는 기상 시나리오는 각각 81개, 87개이며, 이를 통합하여 총 92개의 시나리오로부터 수확량 분포를 도출하였다. 이에 따른 수확량 예측 결과는 건조중량 기준 평균 1.33 ton/ha , 표준편차 0.25 ton/ha 였으며, 수확량 분포에서 $1.2\text{--}1.4 \text{ ton/ha}$ 가 가장 높은 것으로 나타났다.

해당 알고리즘을 적용함으로써 동일 작물 내에서도 당해 기상 조건, 작물 품종 등에 따른 생육 차이를 반영한 수확량 예측이 가능할 것으로 사료되며, 스마트 농업단지에서 다양한 데이터가 축적되면 모형 보정을 통해 더 정확한 수확량 예측 모형으로 활용될 수 있을 것으로 기대된다.

6. 마치면서

본 고에서는 작물 생육 모형과 기상 시나리오를 결합하여, 작물 생육 상태를 반영한 수확량 예측 알고리즘을 제안하였다. 특히 수확량을 단일 값이 아닌 분포 형태로 제시하고, 특정 시점의 생육

상태를 기반으로 유사 시나리오를 선별하여 예측의 현실성을 높이고자 하였다. 이는 스마트 농업단지에서 측정된 작물 생육 데이터를 수확량 예측에 활용할 수 있는 하나의 접근 방법으로서 의미가 있으며, 자료가 제한된 환경에서도 모형 기반 모의를 통한 수확량 예측의 가능성을 제시한다.

다만 본 연구는 실제 수확량 데이터가 부족한 상황에서 수행되었기 때문에 모형 보정이 충분히 이루어지지 못했다는 한계가 있다. 또한 적용된 작물 모수는 특정 지역에서 보정된 값으로, 대상 지역의 품종 특성을 충분히 반영하지 못할 가능성도 존재한다.

그럼에도 향후 스마트 농업단지에서 다양한 생육 및 재배 데이터가 축적될 경우, 이를 활용한 모형 보정을 통해 예측 정확도를 향상시킬 수 있을 것으로 예상된다. 아울러 해당 알고리즘은 재배관리 정보와 실시간 생육 데이터를 반영한 정밀 예측 체계로 확장될 수 있으며, 이는 기존의 경험적 접근에서 정량적·예측 중심의 데이터 기반 의사결정 체계로의 전환에 기여할 것으로 기대된다.

참고문헌

1. 농사로, <https://www.nongsaro.go.kr>
2. 농촌진흥청, 2014, 농업기술길잡이 174:귀농귀촌.
3. 농촌진흥청, 2021, 농업기술길잡이 116:공.
4. 윤푸른, 2022, 기후변화에 따른 시설재배와 논의 물-에너지-식량-탄소 넥서스 분석, 서울대학교 대학원 박사학위논문.
5. 윤푸른, 김마가, 최진용, 2025, 농업환경보전프로그램 농업 재배활동의 넥서스 기반 영향.
6. 최순근, 김민경, 정재학, 최동호, 허승오, 2017, APEX-Paddy 모델을 이용한 기후변화에 따른 논벼 생산량 및 증발산량 변화 예측, 한국농공학회 논문집, 59(4): 27-42.
7. 통계청, <https://kosis.kr>
8. 허승오, 최순근, 홍성창, 2019, APEX 모형을 이용한 밭작물(공, 배추) 물발자국 영향 평가, 한국환경농학회지, 38(3): 159-165.
9. 흙토람, <https://soil.rda.go.kr>
10. Steglich, E. M., J. Osorio, L. Doro, J. Jeong, and J. W. Williams, 2023, Agricultural Policy/Environmental Extender Model User's Manual Version 1501, Blackland Research and Extension Center.
11. World Economic Forum, 2016, The future of jobs: Employment, skills and workforce strategy for the fourth industrial revolution.
12. Williams, J. R., R. C. Izaurralde, C. Williams, and E. M. Steglich, 2023, Agricultural Policy/Environmental Extender Model Theoretical Documentation Version 1501, Blackland Research and Extension Center.